COMMITTEE CERTIFICATION OF APPROVED VERSION

The committee for Charles Todd Davis certifies that this is the approved version of the following dissertation:

MOLECULAR EPIDEMIOLOGY OF WEST NILE VIRUS IN NORTH AMERICA

Committee:

Alan D.T. Barrett, Ph.D.

Supervisor

Nigel Bourne, Ph.D.

T. Kley Hughes, Ph.D.

Robert S. Lanciotti, Ph.D.

Robert B. Tesh, M.D.

Dean, Graduate School

MOLECULAR EPIDEMIOLOGY OF WEST NILE VIRUS IN NORTH AMERICA

By

Charles Todd Davis, B.S., M.S.P.H.

Dissertation

Presented to the Faculty of

The University of Texas Graduate School of
Biomedical Sciences at Galveston

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Approved by the Supervisory Committee:

Alan D.T. Barrett, Ph.D.
T. Kley Hughes, Ph.D.
Robert S. Lanciotti, Ph.D.
Nigel Bourne, Ph.D.
Robert B. Tesh, M.D.

September, 2005
Galveston, Texas

Key words: Flavivirus, phylogenetics, genetic/phenotypic variation, molecular epidemiology

To my wife Lauren

ACKNOWLEDGEMENTS

I would like to begin by thanking my mentor, Dr. Alan Barrett, for his guidance throughout my four years at the University of Texas Medical Branch. His overwhelming energy and tireless work ethic have not only made the past four years extremely productive, they have made the experience of working in his laboratory stimulating and enjoyable. He will continue to be a role model throughout my career as a scientist and in my personal life. I would also like to offer my sincere thanks and gratitude to Dr. David Beasley who was instrumental in my training as a bench scientist and who spent countless hours helping me to learn the many technical aspects of the research presented in this dissertation. I would also like to thank Dr. Scott Weaver, Dr. Doug Watts, Dr. Robert Shope, Dr. Norbert Herzog, and Dr. Steve Higgs for their help and advice over the course of this dissertation. I am indebted to Harriet Ross for her aid as the assistant to the Experimental Pathology Graduate Program. I must thank Dr. Greg Ebel and Dr. Aaron Brault for their kind contribution of West Nile virus isolates and sequence data used in this dissertation and for their helpful discussions concerning the findings of my research. I also want to thank all of the members of the Barrett laboratory, both past and present, including Dr. Mike Holbrook, Dr. Tom Solomon, Dr. Juliet Bryant, Amber Engel, Jason Wicker, Melissa Whiteman, Greg Gromowski, Dr. Jana von Lindern, Dr. Monica McArthur, Li Li , and Shuliu Zhang for all of their help throughout the past three years. I am especially grateful to Li Li and Shuliu Zhang for their tireless contributions to many of the technical aspects of this dissertation.

I must thank members of my dissertation committee, Dr. Robert Lanciotti, Dr. Kley Hughes, Dr. Nigel Bourne, and Dr. Robert Tesh for taking the time from their busy schedules to oversee all of this research. I am especially grateful to Dr. Tesh and members of his laboratory for their support, without which this dissertation would not have been possible. I am also grateful to Dr. Robert Lanciotti for his willingness to contribute sequence data for this project and for his contributions as a collaborator throughout this process. I want to thank Brad Schneider, Lark Coffey and Yvette Girard for their willingness to participate in the many extracurricular activities available in and around Galveston Island. I would also like to thank Dr. Peter Mason for his eagerness to offer constructive ideas for my research and for his willingness to lend a hand and hammer during my home “restoration” project. I would also like to thank the UTMB Protein Chemistry Laboratory for nucleic acid sequencing. The funding for this research was provided in part by the State of Texas Advanced Research Program, by the James W. McLaughlin Fellowship Fund, and by an NIH T32 Training Grant for the research of Emerging Infectious Diseases (T32 AI 007536). I am also indebted to my parents for their contributions over the years both personally and financially. My many years of schooling would not have been possible otherwise. Finally, I am forever grateful to my wife, Lauren, for her patience, determination, and positive attitude even in the face of “temp” work in downtown Galveston. Most of all, I am grateful for her love and kindness as my wife.

MOLECULAR EPIDEMIOLOGY OF WEST NILE VIRUS IN NORTH AMERICA

Publication No._____

Charles Todd Davis, B.S., M.S.P.H.

The University of Texas Graduate School of Biomedical Sciences at Galveston, 2005

Supervisor: Dr. Alan D.T. Barrett

The introduction of West Nile virus (WNV) into the U.S. during the summer of 1999 led to the largest epidemic of arboviral encephalitis ever recorded in the Western Hemisphere during 2002. Over the course of six years, the distribution of WNV has expanded to include each of the contiguous U.S. states and seven Canadian provinces, as well as Mexico, several Caribbean Islands, and Colombia. In order to understand how the virus has evolved since its emergence in North America, this dissertation investigates the genetic and phenotypic variation among WNV isolates collected in various regions of North America during different transmission seasons. The overall objectives of this dissertation were to study the extent to which WNV has evolved since its emergence in North America and to better understand the relationship between viral evolution and phenotypic variation in an emerging viral population. The first aim of this project was to compare nucleotide and deduced amino acid sequences of WNV isolates collected in North America during 2002, 2003, and 2004 to those collected during earlier years. Sequence comparisons of WNV isolates collected throughout North America during

different years have identified nucleotide/amino acid substitutions that reveal the emergence of genetically divergent variants of WNV in recent years and support the hypothesis that microevolution of WNV will continue from year to year and as the distribution of the virus expands. In addition, these studies have identified a dominant genotype of North American WNV that has displaced all other known genotypes throughout North America. This research also led to the discovery of several genetic variants with altered phenotypes. Thus, the second aim of this project investigated the phenotypic characteristics of WNV variants collected in 2003 and 2004 by evaluating their small plaque and temperature-sensitive phenotypes, their multiplication kinetics in cell culture, and attenuation in a mouse model. Finally, reverse genetic techniques were used in a third aim in order to precisely identify the mutations responsible for the observed phenotypic changes in WNV and to illustrate novel molecular mechanisms of attenuation of WNV. The results from these aims demonstrate that as WNV has accumulated mutations in its genome, phenotypic variants have emerged with significantly different biological properties when compared to progenitor virus isolates that initiated the epidemic in North America. Also, this study has identified novel molecular determinants of attenuation in WNV that provide valuable insight into the multigenic components of pathogenicity for WNV and possibly other related flaviviruses.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

µl	microliter
3’ UTR	3’ untranslated region
5’ UTR	5’ untranslated region
Arg (R)	arginine
Asn (N)	asparagine
Asp (D)	aspartate
AST	average survival time
C protein	capsid protein
CDC	Centers for Disease Control
CSF	cerebral spinal fluid
CNS	Central Nervous System
CPE	cytopathic effect
Cys (C)	cysteine
DEN	dengue
DHF	dengue hemorrhagic fever
dNTPs	deoxyribonucleotide triphosphate
DTT	dithiothreitol
EEE	Eastern equine encephalitis
E protein	envelope protein
ER	endoplasmic reticulum
Gln (Q)	glutamine
Glu (E)	glutamate
Gly (G)	glycine
HCV	hepatitis C virus
His (H)	histidine
HPLC	high performance liquid chromatography
Ile (I)	isoleucine
ip	intraperitoneal
ic	intracerebral
JE	Japanese encephalitis
kb	kilobase
kDa	kilodalton
LD_{50}	lethal dose 50
Leu (L)	leucine
Lys (K)	lysine
M protein	membrane protein
m.o.i.	multiplicity of infection
Met (M)	methionine
ml	milliliters
mRNA	messenger RNA

NAMRU	Naval Medical Research Unit
ntd.	nucleotide
NS1 protein	nonstructural 1 protein
NS2A protein	nonstructural 2A protein
NS2B protein	nonstructural 2B protein
NS3 protein	nonstructural 3 protein
NS4A protein	nonstructural 4A protein
NS4B protein	nonstructural 4B protein
NS5 protein	nonstructural 5 protein
PCR	polymerase chain reaction
PD_{50}	protective dose 50
PFU	plaque forming unit
Phe (F)	phenylalanine
prM protein	premembrane protein
Pro (P)	proline
RdRp	RNA-dependent RNA polymerase
RNA	ribonucleic acid
RT	reverse transcription
Ser (S)	serine
SLE	St. Louis encephalitis
TBE	tick-borne encephalitis
Thr (T)	threonine
Trp (W)	tyrptophan
Tyr (Y)	tyrosine
UTMB	University of Texas Medical Branch
Val (V)	valine
Vero	African green monkey kidney cells
VSV	vesicular stomatitis virus
WNV	West Nile virus
YF	Yellow fever

CHAPTER 1

INTRODUCTION

Natural history of West Nile virus

Overview

Historically, human infections with West Nile virus (WNV), a member of the genus *Flavivirus*, family Flaviviridae, were associated with a mild undifferentiated fever. However, recent outbreaks in Europe, Israel, and North America involving humans, equines, and birds have been associated with significant rates of neurological disease (Lanciotti et al., 1999; Solomon and Vaughn, 2002). The virus is maintained in nature in a natural transmission cycle involving many species of birds that act as reservoir and amplifying hosts for the virus and the mosquito vector. The virus is primarily transmitted from bird to bird by ornithophagous mosquitoes, such as members of the *Culex* genus, but has been isolated from a wide variety of mosquito species belonging to numerous genera. Until 1999, the geographical distribution of the virus was limited to Africa, the Middle East, India, Australasia, and western and central Asia with occasional epidemics occurring in Europe (Murgue et al**.,** 2002). As of 2005, however, the distribution of the virus has expanded to include the entire continental United States (U.S.) and many Canadian provinces, as well as Mexico, a number of Caribbean Islands, and parts of

Central America, with the most recent reports coming from Columbia (Kramer, personal communication) (Figure 1-1). Due to this rapid expansion, WNV now has the largest geographic distribution of all the members of the *Flavivirus* genus.

Classification of West Nile virus

West Nile virus was first isolated in the West Nile district of Uganda in 1937 from the blood of a viremic woman participating in a yellow fever study (Smithburn et al., 1940). Since its discovery as a novel virus entity, WNV has been identified as the causative agent of both epizootics and epidemics throughout Africa, the Middle East, Europe, Asia, Australasia (Kunjin virus), and North America (Murgue et al., 2002). The earliest reports of epidemics due to WNV came from Israel and Egypt in the 1950's and led to some of the first studies of the seroprevalence and ecology of this virus (Taylor et al., 1956). The resulting studies revealed that WNV was transmitted from mosquitoes to birds, like many previously identified arboviruses, and was capable of infecting and causing illness in humans and horses. West Nile virus was first classified serologically as a member of the *Flavivirus* genus due to its antigenic relatedness to members of this group. Cross-neutralization studies then revealed that WNV was most closely related to particular flaviviruses, such as Japanese encephalitis (JE), St. Louis encephalitis (SLE), Murray Valley encephalitis, and Usutu viruses, and was subsequently classified as a member of the JE serocomplex (Smithburn, 1942; Hayes, 2001b; De Madrid, 1974). Hemagglutination-inhibition studies showed that strains of WNV isolated from different countries demonstrated antigenic variation and that these isolates could be divided into

subgroups, or lineages (Hammam and Price, 1966). Early serological studies demonstrated the relationships of WNV isolates used polyclonal antisera and monoclonal antibodies and identified three geographically distinct antigenic complexes: viruses predominantly from Africa, viruses from India, and viruses predominantly from Europe and the Middle East (Blackburn et al., 1987). With the advent of nucleic acid sequencing, additional studies have since confirmed the placement of geographically distinct virus isolates into the same lineages that were originally proposed by serological assays (Berthet et al., 1997) (Figure 1-2).

West Nile virus phylogenetics

Subsequent to the studies of Berthet et al. (1997), phylogenetic studies have classified WNV isolates into two major lineages and several subclades (Jia et al., 1999; Lanciotti et al., 1999; Lanciotti et al., 2002; Scherret et al., 2001; Beasley et al., 2002). WNV isolates from Lineage 1 have been isolated from West Africa, the Middle East, Eastern Europe, and North America, while Lineage 2 contains virus isolates from Africa and Madagascar only, with the exception of a few viruses isolated from viremic migratory birds in Mediterranean breeding grounds. Interestingly, Kunjin virus (an Australian relative of WNV) has been found to be more closely related genetically to lineage 1 viruses and is consistently placed in a basal subclade within this lineage upon phylogenetic analysis. Also, of interest is the positioning of Indian WNV isolates in the most basal subclade of lineage 1 (Figure 1-2). This finding is noteworthy in that early hemagglutination-inhibition studies and cross-neutralization tests showed that Indian

isolates did not cross-react strongly with many viruses in the lineage 1 clade (Hammam et al., 1965). Because no complete genomes of Indian viruses have been sequenced, it remains unclear whether or not Indian viruses should be included in subclades of lineage 1 or if they in fact make up a separate lineage. Additionally, bootstrap support of this subclade is weak in comparison to other subclades within lineage 1. Further analysis will be required to explore this possibility. Also, several recent publications have suggested that virus isolates from the Czech Republic (Rabensburg isolate, GenBank accession no. AY765264) and the Caucasus region of Russia (LEIV-Krnd88-190, GenBank accession no. AY277251) may comprise new lineages of WNV (proposed lineages 3 and 4) (Bakonyi et al., 2005). Clearly, additional nucleotide sequence information and serological studies are needed to address this issue and to determine if the classification of WNV into only two lineages should be re-evaluated.

Previous to the introduction of WNV into North America, it was perceived that Lineage 2 WN viruses were the only viruses consistently found in areas with an endemic enzootic transmission cycle and where serological surveys have revealed high levels of anti-WNV antibodies in the human population (Morvan et al., 1990; Omilabu et al., 1990). Lineage 1 viruses, on the other hand, have been isolated from regions of non-endemic foci, where humans, birds, and equines presumably have less protective immunity. These viruses have also recently been associated with epidemics of encephalitis in Europe, Russia, Romania, Israel, and the U.S., while viruses from lineage 2 have not been associated with recent outbreaks (Scherret et al., 2001). These observations, together with the recent introduction of a highly virulent lineage 1 WNV

into North America, have led researchers to investigate the genotypic differences among members of lineage 1 and 2 WN viruses that correlate with differences in virulence phenotypes. A recent study by Beasley et al. (2002) revealed the existence of several strains of both lineages 1 and 2 that are naturally attenuated in neuroinvasiveness as demonstrated by mouse and hamster models. Thus, it is clear from these studies that both virulent and avirulent strains of WNV from lineage 1 and 2 can be found in naturally occurring WNV transmission cycles, further obscuring the relationship between epidemic and endemic cycles of WNV and their impact on human and veterinary health.

Biology and ecology

Early studies conducted by the United States Naval Medical Research Unit (NAMRU) in the Nile Delta concerning the ecology of WNV revealed that this virus was transmitted via mosquito vectors to birds, which acted as reservoir and amplifying hosts. Serological studies demonstrated that humans and horses could be infected, but that infection normally resulted in only a mild febrile disease (Taylor et al., 1956). Experiments in both humans and equines later suggested that horses and humans were most likely dead-end hosts of the transmission cycle because serum viremia titers following peripheral infection were low, even though high titers of neutralizing antibodies were produced (Hayes, 2001; Southam and Moore et al., 1954). Results from these early studies suggested that WNV resulted in a self-limiting febrile illness, primarily infecting children early in life that was rarely associated with encephalitis. Serological studies of other mammals revealed that a large number of animals, primarily

quadrupeds, could also be infected by WNV and that many species of birds could serve as reservoir and amplifying hosts (Taylor, 1956). Many early serological surveys of antibody prevalence in humans were conducted in the Nile Delta region and showed a high prevalence, particularly in children (73.9% in northern Egypt) (Darwish & Ibrahim, 1975;Hayes, 2001a).

Episodic epidemics and epizootics of WNV have occurred since the discovery of WNV, and it was perceived that these periodic outbreaks were depended on climatic conditions that allowed for a large number of mosquitoes to be in contact with numerous bird reservoirs in close proximity to human and/or equine populations. Thus, early investigators considered WNV to be a disease agent of tropical and subtropical climates where mosquito populations were sustained by warm temperatures and heavy rainfall and where bird populations were abundant. Epidemics in colder climates, such as those in Eastern Europe, were believed to have initiated in summer months when migratory birds carried viruses into areas with large mosquito populations (Hayes, 2001). Thus, epidemics were believed to occur in waves that would die out because of a failure of the virus to establish an endemic transmission cycle. This narrow view of the ecology of WNV was quickly dispelled when the virus was introduced into the northeastern U. S. and was able to persist through the cold winters and spread rapidly despite the temperate climate. Several factors have been proposed to explain how WNV has been able to maintain an endemic cycle in colder climates. Evidence of virus isolated from birds in winter months suggests that perhaps some species of birds are able to be persistently infected and that mosquitoes may pick up the virus when temperatures increase (Komar,

2002). This is not considered likely, as viremias in birds would need to be high enough to allow mosquito infection without resulting in disease to the bird during winter months. Studies have also shown that transovarial transmission of virus from adult to progeny may provide a source of infected mosquitoes from one season to the next (Miller et al., 2000). Perhaps more feasible as a means of sustained virus circulation is the ability of infected mosquitoes to overwinter, even in harsh climates where temperatures fall below freezing, and initiate transmission when temperatures rise. Additionally, there is evidence to show that year-long transmission is sustained in parts of the U. S. where winter temperatures are high enough to allow feeding behavior in the mosquito all year (Tesh et al., 2004). Migratory birds would then be capable of re-introducing virus into cold weather regions during their annual migrations. Regardless of the mechanism(s), WNV was able to establish endemic transmission in many temperate regions of the U.S. and Canada and continues to expand its range in North America every year. It is now clear that many factors other than climate, vectors, and reservoir hosts contribute to the overall transmission and maintenance of the virus in nature (Figure 1-3).

Epidemiology

Despite the high prevalence of WNV antibodies in regions of the Nile Delta in the late 1960's, epidemics of disease were rarely reported prior to introduction of the virus into the U.S. The first reported outbreak occurred in Israel in 1951-1952 with over 50% of the cases found in children less than 6 years of age (Murgue et al., 2002). Since then, disease outbreaks have been reported at various times from Israel (1957 and 1962),

France (1962, 2000), South Africa (1974, 1983-1984), Algeria (1994), Romania (1996), Morocco (1996), Tunisia (1997), Italy (1998), Russia (1999), and the United States and Canada (1999-present) (Murgue et al., 2002) with sporadic disease reported from many African, Middle Eastern, and European countries. Prior to introduction of the virus into the U.S., annual world-wide case numbers were in the hundreds with cases of encephalitic disease reported very rarely. As with other diseases common to developing countries, there are likely to be a number of cases that go underreported every year. However, it is also likely that low numbers are reported because infection with WNV most commonly results in asymptomatic infection. Epizootic disease outbreaks, involving birds and horses, have also been scarcely reported, but the number of these reports have steadily increased since the discovery of WNV in the U.S. (Hayes, 2001c). Epidemics have commonly been associated with climate conditions that permit large increases in local mosquito populations (high temperatures and heavy rainfall) allowing for increased contact between mosquitoes and humans, horses, and birds. Thus, it is very likely that epidemics will continue to occur in regions of the world when the right conditions prevail.

Disease manifestations

WNV infection in humans typically results in either asymptomatic infection or mild illness with fever. During some epidemics, reports of an acute, febrile illness with mild to severe muscle weakness, lymphadenopathy, backache, retro-orbital pain, and rash have also been described. Although the majority of cases found during outbreaks of

WNV are classified as having a mild, febrile disease, neurological symptoms have been reported as well. The first reported cases of severe neurological symptoms were in Israel in 1957 and have since been described many times in various geographical regions (Solomon and Vaughn, 2002). Encephalitic disease caused by WNV can range from mild to severe meningoencephalitis, flaccid and long-term paralysis, as well as diffuse encephalitis and death (Hayes et al., 2005a). Pathological findings reveal that damage to the central nervous system (CNS) is similar to that seen in JE or poliomyelitis, with inflammatory infiltrates in the brain and spinal cord, perivascular cuffing, and the presence of glial nodules (Solomon and Vaughn, 2002). While encephalitis remains a rare complication of WNV infection, it is believed that the elderly, organ transplant recipients, and other immunocompromised people are at greater risk of the manifestation of neurological disease (Solomon and Vaughn, 2002; Granwehr et al., 2004). It has also been hypothesized that diabetes, hypertension, and cerebrovascular disease are risk factors, as well, although this remains to be confirmed (Granwehr et al., 2004). Neuroinvasive disease is thought to develop in <1% of infected persons, and it should be pointed out that the proportion of reported cases that show neuroinvasive disease is likely to be high because neuroinvasive illness is more readily reported than asymptomatic infections (Hayes et al., 2005a).

Pathogenesis

WNV is believed to replicate at the site of inoculation by the mosquito, most likely by infecting dendritic cells, and then multiplies in tissues and lymph nodes near the

site of entry. The virus then moves to the blood via the lymphatics and is capable of reaching many organ systems including the heart, liver, spleen, pancreas, kidney, intestine, eye, and CNS. Viral penetration of the CNS has been proposed to occur following stimulation of toll-like receptors (i.e., TLR-3) and increasing levels of tumor necrosis factor-α, which are thought to lead to increases in vascular permeability at the blood-brain barrier (Wang et al., 2004). It has also been postulated that WNV may enter the CNS by invasion of the peripheral nerves or by retrograde transport via the olfactory nerves (Monath, 2002). WNV is highly neurotropic once inside the CNS and readily infects neurons, especially in deep nuclei and gray matter of the brain, brainstem, and spinal cord. Destruction of bystander nerve cells may also contribute to neurological symptoms associated with WNV induced CNS disease. Immune-mediated nervous tissue damage has also been suggested to result in CNS disease, as evident by the presence of inflammatory infiltrates upon histological examination (Hayes et al., 2005). Most recent studies concerning the pathology of WNV, and especially postmortem studies, have concerned infection with the North American strain of WNV. However, as noted in earlier sections, various strains of WNV exhibit significantly different virulence phenotypes that are no doubt related to molecular differences at the genome level. Therefore, it is likely that the pathological findings present here vary considerably depending on the strain of virus in question. Clearly, the WNV strain introduced into North America possesses molecular properties conferred by its genome that are responsible for its highly virulent phenotype.

Molecular Virology

Genomic organization

The genome of WNV is a single-stranded, positive-sense RNA molecule comprised of a single open reading frame (ORF) that is flanked on both ends by a 5' and 3'untranslated region (UTR). The ORF is translated into a single polyprotein that is both co- and post-translationally cleaved to produce three structural proteins (capsid, premembrane, and envelope) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5) (Figure 1-4). The genome is approximately 11,000 nucleotides in length and is capped at the 5' end by a type 1 m^7GpppAmp structure and terminates at the 3' end with a CU_{OH}. The WNV genome serves as the only mRNA during translation of viral proteins. Host signal peptidases are thought to cleave the C-prM, prM-E, E-NS1, and C-terminus of NS4A. A virus encoded serine protease is believed to cleave each of the nonstructural proteins with the exception of NS1-NS2A, whose cleavage enzyme remains unknown (Lindenbach and Rice, 2003). The following section will briefly touch on the structural and functional roles of each of the WNV proteins and untranslated regions.

WNV structural proteins

<u>Capsid (C) protein</u>

The C protein gene encodes a highy basic 11 kDa protein that forms the viral nucleocapsid surrounding the viral RNA. The protein contains a short C-terminal

hydrophobic domain that serves as a signal peptide for ER translocation of the premembrane protein. This domain is cleaved from the mature C by the viral serine protease (Amberg and Rice, 1999). Charged residues localized around the N- and C-terminus of the protein are believed to mediate the interaction with viral RNA (Khromykh and Westaway, 1996). Recently, the solution structure of the DEN-2 C protein was solved revealing that the C protein dimer has an asymmetric distribution of basic residues over the surface and a conserved hydrophobic region, which are proposed to enable the interaction of the C protein with viral RNA and the membrane protein (Ma, 2004).

Premembrane (prM)

The prM protein has been shown to function in maturation, stabilization, and proper folding of the envelope protein during formation of the virion. The prM is a 26 kDa protein that is a precursor to the viral membrane (M) protein and is cleaved by host furin-like proteases before the release of the virion from the cell. This cleavage event has been shown to coincide with the conversion of immature virus particles to mature virions. Analysis has demonstrated that prM and E form a heterodimer after synthesis and the prM has a chaperonin-like activity that is involved in maintaining the folding pattern of the E protein. Additional studies have shown that the pr segment of the prM stabilizes the E protein and keeps it from rearranging into its fusogenic form (Lindenbach and Rice, 2003).

Envelope (E) protein

The WNV E protein is the major viral surface protein mediating cellular binding and membrane fusion. The E protein is located at the surface of the virion as a series of head to tail homodimers rich in β sheets with transmembrane domains anchoring the protein to the membrane of the virion. Much of what is known about the WNV E protein is based on previous studies with other flaviviruses, including tick-borne encephalitis (TBE) virus (Rey et al., 1995) and dengue-2 virus (DEN-2) E proteins (Modis et al., 2003). These studies have shown the flavivirus E protein to be a class II fusion protein with each monomer made up of three domains (I, II, and III). During fusion of the virion with host cell membranes, a hinge region between domains I and II allows the protein to form trimers that expose the fusion peptide at the tip of domain II (a pH-dependent conformational reaction). Prior to fusion, the E protein dimers lie parallel to the surface of the virion with residues important for receptor-binding located on the outermost surface. Domain III contains the putative receptor-binding domain that is thought to function in the binding of the virion to host cell receptors. The immunoglobulin-like folding pattern and RGD integrin-like sequence of DIII, as well as a number of receptor-binding studies and mutational analyses, support the hypothesis that this domain is responsible for cell surface interactions between the virus and host (Heinz and Allison, 2003) (Figure 1-5).

Because of its external position surrounding the virion, the E protein is also the primary immunogen of WNV and often the focus of rationale vaccine design (to be covered in later sections). The E protein of WNV may or may not contain a single N-

linked glycosylation site that becomes glycosylated by oligosaccharides during infection. Recent studies have shown that variation in the glycosylation status of the WNV E protein can lead to changes in the pathogenicity of the virus perhaps by stabilizing the protein or by protecting it from pH induced degradation (Beasley et al., 2005; Shirato et al., 2004). In these experiments, strains possessing the E protein glycosylation motif have increased neuroinvasiveness in mice following peripheral inoculation when compared to strains that lack this motif. It has been proposed that the presence of an E protein glycosylation motif is one of the reasons why the North American strain of WNV is so highly virulent in birds and has the potential to induce encephalitis in horses and humans.

WNV nonstructural proteins

NS1 protein

NS1 is a glycoprotein that has up to three potential N-linked glycosylation sites, whose glycosylation status has also been hypothesized to be a determinant of WNV pathogenicity (Whiteman, unpublished data). Although the function of NS1 remains largely unknown, several studies of flavivirus NS1 mutants have suggested that this protein plays a role in RNA synthesis. NS1 translocates to the ER prior to proteolytic processing and has been shown to be secreted by mammalian cells (Brinton, 2002). Other studies have shown that NS1 is associated with intracellular membranes during cellular infection, but that the protein does not directly associate with the mature virion. In

addition, trans-complementation studies have also shown that this protein is required for both positive and negative strand RNA synthesis (Lindenbach and Rice, 2003).

Small, hydrophobic nonstructural proteins – NS2A, NS2B, NS4A, and NS4B

Each of these four proteins is a hydrophobic, membrane-associated protein. None of their sequence motifs share homology with any currently known enzymes. It has been proposed that they each play some role in assembly of the replication complex by binding to other host or viral proteins. NS2A cleavage from NS1 appears to be a requirement for NS1 processing and has been shown to associate with NS4A to allow for RNA packaging in the ER (Brinton, 2002). NS2B complexes with and serves as a co-factor for NS3 through a hydrophilic domain (Chu and Westaway, 1985), while a hydrophobic domain, on the other hand, may facilitate cotranslational insertion of the NS2B/NS3 precursor into the ER. Perhaps the least understood of the flavivirus proteins, NS4A has been found to colocalize with NS1 and the viral 3'UTR, and has been hypothesized to play a role in replication complex stabilization (Lindenbach and Rice, 1999). The function of NS4B is also poorly understood. However, recent interest in its role as a potential inhibitor of the interferon-signaling cascade during DEN infection has sparked interest in its function (Munoz-Jordan et al., 2005). Studies have described the accumulation of WNV NS4B protein in the perinuclear region of infected cells and the ability of NS4B to translocate to the nucleus (Westaway et al., 1997). Other research has shown NS4B mutations in DEN-4 virus that influence mosquito infectivity and the susceptibility of SCID mice to viral infection (Hanley et al., 2003).

NS3

NS3 is known to possess several enzymatic functions. This protein functions as the viral serine protease, helicase, nucleoside triphosphatase (NTPase), and RNA triphosphatase (RTPase). The serine protease activity is encoded in the N-terminal region and is critical for cleavage of the viral nonstrucutral proteins following translation of the polyprotein. NS3 and NS5 form a major part of the replicase complex and have been found to colocalize in the perinuclear region. The C-terminus of NS3 contains regions of homology with motifs expressing helicase, NTPase, and RTPase activity. It has been postulated that the helicase, together with the NTPase, function in the unwinding of viral RNA. The RTPase is believed to function in the synthesis of 5'cap formation (Brinton, 2002).

NS5

The largest and most conserved of the flavivirus proteins, NS5 functions as both a methyltransferase encoded by the N-terminal region, required for 5'capping, and as the viral RNA-dependent RNA polymerase (RdRp), essential for replication. Phosphorylation of NS5 at serine residues has been shown to be necessary for polymerase activity and in order for the NS3/NS5 complex to form (Kapoor et al., 1995). The NS5 of DEN-2 virus has also been shown to contain a nuclear localization signal that may direct the protein or its complexes to the nucleus of the host cell (Lindenbach and

Rice, 2003). Mutational studies have shown that both the methyltransferase and the RdRp are required for viral replication (Khromykh et al., 1999).

WNV untranslated regions

5'UTR and 3'UTR

The 5'UTR and 3'UTR of WNV are approximately 100 and 600 nucleotides in length, respectively, although the length of the 3'UTR varies considerably between different strains. Both the 5' and 3' UTR are important for viral translation and as the initiation sites of RNA synthesis during replication. Each possesses conserved secondary structures (stemloops, pseudoknots, and dumbells) that are critical for replication. These structures have also been hypothesized to play a role in protecting the RNA from digestion by cellular nucleases. Additional studies have demonstrated that the stemloop structures are important for the formation of binding sites for cellular or viral proteins. The 5'UTR and the 3'UTR are thought to come together during translation initiation as a closed-loop complex. This is believed to occur through either interactions with the 5'cap structure and cellular or viral proteins or by base pairing interactions between 3' and 5' cyclization sequences (Markoff, 2003). In any case, formation of this complex is an essential step in the translation of viral proteins.

Viral life cycle

The virus life cycle of WNV is similar to other positive-sense, enveloped viruses. The virion attaches to the surface of the host cell and enters the cell by receptor-mediated endocytosis most likely through interactions between DIII of the E protein and a yet to be determined host cell receptor(s) (Figure 1-5). Acidification and the accompanying decrease in pH of the endosomal vesicle triggers conformational changes in the E protein dimer as described previously. Fusion of the viral and cell membrane then occurs, resulting in virion disassembly, release of the nucleocapsid into the cytoplasm, and dissociation of the capsid from viral RNA. After the genomic RNA has been released into the cytoplasm, it is translated as a long polyprotein that is then cleaved co- and post-translationally by viral and host proteases into the structural and nonstructural proteins. Replication of the new genomes occurs on intracellular membranes near the endoplasmic reticulum (ER) and immature virion assembly occurs at the ER surface (Figure 1-6). Structural proteins and genomic RNA are believed to assemble and bud into the ER lumen, giving rise to immature viral particles that are transported through the Golgi towards the host cell membrane. Immature virus particles are cleaved by host cell furins in the trans-golgi network (TGN) and mature virus is then released by exocytosis to the outside of the cell (Mukhopadhyay et al., 2005).

Replication

As stated above, replication of the WNV genome occurs on intracellular membranes. The viral RdRp, NS5, functions by synthesizing complementary negative-strands by transcription of the positive-strand genomic RNA. Newly synthesized negative-strand RNAs then act as templates for the synthesis of new positive-strand genomic RNAs. Several studies have shown that positive-sense RNA synthesis is roughly 10 times more efficient than negative-strand RNA synthesis. Once established, both positive and negative-strand viral RNA synthesis is likely to continue after protein synthesis has abated, indicating that viral polyprotein is not a constant requirement for RNA synthesis. Interestingly, studies suggest that nascent genomic RNA can function as a template for translation, transcription, or as source of virion RNA, although there is evidence that in order for positive-strand RNA to be used as a replication template, translation must have begun, and that encapsidation of the viral RNA does not happen until replication has occurred. Early during the replication cycle, nascent RNA may be used as a template for both replication and translation until enough structural proteins have been translated for virion assembly (Brinton, 2002).

Emergence of West Nile virus in North America

Introduction, range expansion, and epidemiology in the U.S.

West Nile virus was first detected in New York City during the summer of 1999 when an outbreak of encephalitis in elderly patients led investigators to consider an

arboviral etiology. Initially, St. Louis encephalitis (SLE) virus, which has similar symptoms to WNV and was known to be enzootic on the east coast of the U.S., was considered in the differential diagnosis. Because of the high degree of cross-reactivity between SLE virus and WNV (i.e., both viruses are in the JE serocomplex), serological testing for antibodies to SLE virus produced positive test results. Several weeks prior to the outbreak of encephalitis, however, there had been reports in the New York City area of widespread bird deaths, and although birds are the natural reservoir for SLE virus, it rarely causes avian mortality. Thus, retrospective laboratory analysis ultimately confirmed WNV as the agent responsible for both the human outbreak and the bird deaths. The initial outbreak involved 62 human cases, including 7 fatalities, and 25 equine cases, including 9 deaths. In 1999, presence of the virus was reported in 4 states. Subsequently, the range expansion increased both westward and southward and was detected in 8 additional states in 2000 and 16 more in 2001. Although the geographical spread of the virus increased considerably in those two years, the number of human cases and deaths attributed to WNV did not. In 2000, there were only 21 cases and 2 deaths reported, and in 2001 there were 48 cases and 5 deaths. Because the case numbers in 2000 and 2001 were low, there remained speculation about whether or not the virus would cause a larger epidemic in humans or even if the virus would become endemic in the temperate climates of the northern U.S. However, despite the lack of human disease, the virus continued to be isolated from mosquito pools, dead birds, and was the cause of illness and death in horses. For reasons discussed previously, WNV was able to maintain

an enzootic transmission cycle despite the temperate climate of the region into which it was introduced.

In the year 2002, the significance of the emergence of this virus in North America was realized. During that year, there were at least 4,156 reported human cases, including 284 fatalities, and over 14,000 equine cases and 12,000 bird fatalities. Additionally, the virus had been detected in all but four states of the continental U.S., reaching California in July 2002 (Figure 1-7). This increase in case numbers and deaths represented the largest epidemic of arboviral encephalitis ever recorded in the western hemisphere (CDC, 2002). The following year, WNV was detected in ever contiguous U.S. state and there were a total of 9,858 human cases reported, including 262 fatalities, and at least 4,000 equine cases and 11,000 bird fatalities (CDC, Arbonet). While the number of horse cases dropped considerably, most likely due to the implementation of an inactivated veterinary vaccine, the number of human cases had risen dramatically as the virus became established in new regions. During 2004, the number of cases was less than 2003 with only 2,448 human cases reported and 87 deaths and a continuing decline in equine cases (Figure 1-8, Figure 1-9) (CDC, Arbonet) and at the time of this writing (August 2005) there have been only sporadic cases of WNV across the U.S. (333 cases of illness; 8 deaths). However, it remains to be seen how the transmission season will proceed. Also, there have been very few reports of WNV disease in either humans, horses, or birds in Mexico over the past several years, even though circulation of virus has been occurring since at least 2002. It has been proposed that the absence of disease may be related to pre-existing protective immunity in humans, horses, or birds that have cross-neutralizing

antibodies to heterologous flaviviruses, such as DEN or SLE viruses, that are endemic to Mexico. Experimental evidence from studies that tested the ability of cross-reactive antibodies to JE or SLE viruses to control WNV infection lends support for this hypothesis (Tesh et al., 2002; Elizondo-Quiroga et al., 2005). Alternatively, it has been proposed that the virulence of strains circulating in Mexico has decreased. This is less likely, as studies done in our laboratory have shown that the majority, though not all, Mexican isolates bear the same phenotypic characteristics as those described in U.S. isolates (this dissertation; Beasley et al., 2004). Finally, it is also possible that cases of WNV have gone underreported or misdiagnosed in areas with other endemic flaviviruses. Regardless, it is presumed that the current epidemic will continue to impact human and veterinary health in the coming years, especially as the geographic distribution of the virus continues to expand southward beyond Mexico and Central America, and into regions where little protective immunity exists in humans and other animals.

Progress towards vaccine design

Two veterinary vaccines are currently available for use in equines. One is an inactivated WNV vaccine based on the WN-NY99 strain that has been in use for over two years now and appears to afford a high level of protection and safety (Ng et al., 2003). A more recently licensed vaccine is a recombinant vaccine that utilizes the canarypox virus to express WNV antigens (prM and E proteins). This vaccine has proven both protective and safe in several experimental animal models (Minke et al., 2004). However, because of its recent licensure, there is little evidence demonstrating its

efficacy in the field. A chimeric live, attenuated virus vaccine that functions by incorporating the prM and E genes of WNV into the yellow fever virus (YF) -17D vaccine backbone has been shown to be efficacious in hamsters and non-human primates and is undergoing phase I clinical trials in humans (Arroyo et al., 2004). A second vaccine candidate is also a live, attenuated chimeric vaccine that incorporates the prM and E genes of WNV into the backbone of an attenuated DEN-4 virus. This vaccine has proven effective in providing protective immunity in non-human primates (Pletnev et al., 2003). A DNA vaccine that expresses the WNV prM and E proteins has also been tested in mice, horses, and birds, and while it elicits some protective immunity, only low levels of neutralizing antibodies were detected in mice and horses prior to challenge (Davis et al., 2001). Further clinical field trials of all of these candidate vaccines will be required before reaching FDA approval.

Phylogenetics of North American West Nile virus

Following the identification of WNV in the U.S. during an outbreak of encephalitis in the New York area in 1999, efforts were made to discover the origin of the virus responsible by analyzing the nucleotide sequence of isolates recovered from dead birds. One of the first isolates made came from the brain of a Chilean flamingo that died from WNV infection at the Bronx Zoo (strain 382-99). Initial partial sequencing of the E protein gene revealed that the isolate was closely related (>99% identity at the nucleotide and amino acid level) to an isolate made from the brain of a dead goose in Israel in 1998 (Lanciotti et al., 1999). Subsequent full-length genomic sequencing of both isolates

confirmed this close relationship (99.7% nucleotide identity) and led to the conclusion that the strain introduced into the U.S. was likely to be of Middle Eastern origin and that it was a lineage 1 WNV isolate (Figure 1-10). Due to the highly neuroinvasive phenotype of the North American isolates, speculation about virulence differences between the NY99 strain and Old World strains resulted in experiments that showed differences in pathogenicity in both mouse and bird models (Beasley et al., 2002; Brault et al., 2004). Clearly, the differences in virulence are dependent on the genotype of individual strains and are not simply the result of single point mutations to the viral genome. Future studies will be needed to define the many molecular determinants of WNV pathogenicity. While we may never know how WNV was actually introduced into the U.S., several possibilities exist. The most plausible explanation is that a previously infected bird or mosquito was carried into the U.S. by airplane or boat. It may also be conceivable, but unlikely that an infected human or other vertebrate host with high viremia carried the virus. Similarly, another unlikely, though not impossible mechanism of introduction would be that a storm displaced or trans-Atlantic migratory bird was able to carry the virus across the Atlantic Ocean.

Ecology of WNV in North America

Another important aspect of the introduction of WNV into North America has been the surprisingly large numbers of species of mosquitoes, birds, and mammals found to be infected by the virus. In the U.S., WNV has primarily been transmitted by *Culex* mosquitoes. However, in contrast to other arboviruses, WNV has been found to be

transmitted by many different genera and species of mosquitoes. To date, at least 59 different species of mosquito from North America have been found infected with WNV. And while the vectorial capacity (ability to amplify and transmit virus) of many of these mosquitoes may vary, clearly the ability to infect so many types of mosquitoes has been important for the establishment of the virus in North America. In fact, the primary northern vector of WNV, *Culex pipiens*, has been implicated in the ability of the virus to survive northern winters because they hibernate as adults and are infectious when warmer temperatures return. Additionally, recent work has suggested that hybridization of *Culex pipiens* (believed to be an ornithophilic mosquito) with a closely related anthropophilic species may have contributed to the high number humans, and perhaps other mammals, that have been infected by the virus (Fonseca, 2004). As of 2004, at least 29 species of mammals, including cats, dogs, sheep, llamas, wolves, goats, squirrels, and skunks have been found infected with the virus. Surprisingly, even alligators have been identified as hosts for WNV with viremias potentially high enough to infect a feeding mosquito (Klenk et al., 2004).

Following the establishment of what appeared to be an endemic transmission cycle in the northeastern U.S. by 2000, WNV quickly spread, presumably by migratory and resident birds, into the warmer climates of the southern U.S. (Figure 1-7). Another species of mosquito, *Cx. quinquefasciatus*, which is able to feed all year long, allowed for the rapid establishment of the virus in this region. By 2002, the virus had been detected across most of the U.S. and southern Canada owing to the fact that everywhere the virus traveled, there was an equally suitable mosquito species to act as the primary vector.

Reports of WNV positive mosquito pools in the western U.S. have implicated the importance of *Cx. tarsalis* in maintaining the transmission of the virus (Granwehr et al., 2004).

Another variable in the equation for establishing sustained transmission of WNV in North America is the abundance of birds that have been discovered to be either infected with or killed by WNV. At least 225 species of birds in the U.S. and Canada have been identified to be infected by WNV over the last six years. And like their mosquito counterpart, there are birds in every region of the continent that are able to successfully amplify the virus and serve as suitable reservoirs. Members of the *Corvidae* family (e.g., crows and jays) have been found to be the most susceptible to infection and also appear to have the highest infectivity titers following experimental infection (Komar et al., 2003). Other birds, such as members of the Colombiformes (pigeons) and Anseriformes (ducks) do not have sufficiently high viremias for transmission (Hayes et al., 2005). Thus, both passive and active surveillance efforts have focused on certain species of birds (primarily the corvids) to act as sentinels of WNV activity and have been found to be reliable indicators of virus transmission activity in a given area (Komar, 2001). Birds, of course, have also been found to play important roles in the dispersal of WNV. Both resident birds and migratory birds have no doubt been responsible for moving the virus over short distances from state to state and over long distances from region to region. Like the mosquito vector, the numerous species of birds involved in WNV transmission have certainly contributed to the rapid and sustained spread of WNV in North America.

Novel mechanisms of transmission

The introduction of WNV into the U.S. has also led to the discovery of novel mechanisms by which WNV can be transmitted. In 2002, intrauterine WNV transmission was documented for the first time in a 20-year-old woman whose child was born with chorioretinitis and cerebral tissue damage. Four additional cases have also been detected with no illness reported in the children, except for one that was born prematurely with respiratory distress. Probable WNV transmission through breast milk has also been reported (Hayes et al., 2005). Transmission of WNV through blood transfusion was also first documented during 2002. Following this discovery, screening of blood donations was started on pools of donated blood using experimental nucleic acid amplification tests. During 2003 and 2004 more than 1,000 WNV positive blood donations were identified. So far at least seven cases of WNV due to transfusion have been documented (Pealer et al., 2003). WNV transmission through organ transplantation has also been described and it has been hypothesized that transplant recipients that are immunosuppressed have an increased risk of severe disease (Iwamoto et al., 2003). WNV infection has also been acquired due to occupational exposure in laboratory workers through needle-stick inoculation and possibly through aerosol exposure (CDC, 2002). There has also been speculation about the possibility of aerosol exposure in workers coming into contact with infected animals as suggested by an outbreak of WNV among turkey handlers at a poultry

farm and alligator handlers at a hatchery (Hayes, 2005). More recently, WNV has also been detected in the urine of a patient with WNV disease (Tonry et al., 2005).

Another recent finding is the potential for WNV to be transmitted by nonviremic transmission from mosquito to mosquito. Previously, it was assumed that in order for a naive mosquito to be infected with WNV, it had to take a bloodmeal from a viremic animal with a serum viremia titer high enough to allow infection. However, recent studies by Higgs et al. (2005) have demonstrated that WNV can be transmitted from mosquito to mosquito without the requirement of replicating virus circulating in the serum of the host. The discovery of nonviremic transmission of WNV is especially relevant because it challenges the dogma that mammals, such as humans and horses, which do not normally mount a sufficient viremia for mosquito infection, are dead-end hosts. While the contribution of nonviremic transmission to the natural transmission cycle of WNV remains unclear, it has the potential to result in increased virus transmission rates and the ability of otherwise dead-end, insusceptible, or immune hosts to spread virus (Higgs et al., 2005).

Specific aims

In order to examine the genetic characteristics of WNV following its introduction into North America, and presumably the whole of the Western Hemisphere, the author has undertaken studies to define the molecular epidemiology of WNV as its geographic and temporal distribution expand. While preliminary data show that microevolution of the virus is occurring on a finite scale, more recent studies provide evidence that as mutations in the WNV genome accumulate, characterization of phenotypic variation in some isolates can be accomplished. As will be demonstrated in this dissertation, several WNV isolates made during the 2003 transmission season exhibited small plaque (sp) morphology, temperature sensitivity (ts) and attenuation of neuroinvasiveness in a mouse model. In other flavivirus models, viral strains exhibiting sp, ts, and attenuated phenotypes in animal models are often useful in defining the genetic determinants of attenuation and virulence for humans, and can be essential when deciding which mutations to incorporate in the development of a live, attenuated vaccine (Hanley et al., 2003; Hanley et al., 2002; Blaney et al., 2003a; Puri et al., 1997; Wallner et al., 1996). Accordingly, this dissertation proposes to examine the phenotypic changes occurring in currently circulating variants of WNV, as compared to variants isolated during earlier years of the North American WNV epidemic, and will rely on nucleic acid sequencing to identify those nucleotide/amino acid substitutions that confer these changes. Use of a WNV infectious clone will then allow for the mapping of the genetic determinants of phenotypic variation occurring in WNV. It remains uncertain whether or not mutations in the virus population will influence the epidemiological significance of the virus in terms

of both human and veterinary health. Thus, from an evolutionary perspective, this study will further characterize the genotypic changes occurring in a recently introduced arbovirus into a unique geographic region, and in doing so, will define important changes in the viral genome that influence the phenotypic characteristics of North American WN viruses. Because of the increasing mortality associated with WNV infection in birds, equines, and humans and the current epidemic showing no signs of decline, it is imperative that research pertaining to the molecular epidemiology of WNV be addressed in order to better understand how this virus will affect human and veterinary health in the future. Therefore, the following specific aims are proposed.

Specific aim 1

Compare nucleotide and deduced amino acid sequences of WNV isolates collected in North America during 2002, 2003, and 2004 to those collected since 1999.

Hypothesis: *Sequence analysis of West Nile isolates representing different geographical locations across North America will reveal genetic variation between viruses collected in different locations and during different years.*

Rationale: Sequence comparisons between WNV isolates collected during 1999-2001 have identified nucleotides/amino acids that differed in isolates circulating in North America during 2002. These results provide evidence that genetically divergent variants of WNV have emerged in recent years. In order to complete this aim, sequence comparisons of isolates made in 2003 and 2004 versus 1999-2002 will be done to reveal the extent of genetic variation in currently circulating WNV.

Specific aim 2(a):

Identification of phenotypic variation in WNV isolates circulating in the U.S. during the 2003 and 2004 transmission season.

Hypothesis: *As West Nile virus continues to circulate from year to year, genetic divergence will result in phenotypic variation among viral isolates.*

Rationale: Based on the results from specific aim 1, it was hypothesized that variation in viral genotype would result in alterations of the biological properties of WNV. Evidence of this association is supported by characterization of genetic variants from 2003 that displayed reduced levels of cytopathic effect in cell culture and small plaque (sp) morphology. In order to identify phenotypic variation in North American WNV, isolates collected in 2003 and 2004 will be compared to progenitor WNV isolates by evaluating the production of sp and ts phenotypes, as well as viral multiplication kinetics in a mammalian cell line, and attenuation of neuroinvasiveness/ neurovirulence in an established mouse model. Evaluation of phenotypic variation among WNV isolates circulating in North America will test the hypothesis that as the virus continues to circulate from year to year, genetic divergence will result in phenotypic variation.

Specific aim 2(b):

Examine the mechanism of attenuation of neuroinvasiveness in 2003 WNV isolates.

Hypothesis: *Isolates attenuated for neuroinvasiveness will induce low levels of viremia following infection and the inability of the virus to invade the central nervous system.*

Rationale: A precise LD_{50} will be determined for 2003 WNV isolates that exhibit a $\geq$1,000 pfu LD_{50} in order to establish the extent of attenuation of neuroinvasiveness in a Swiss Webster mouse model. This model will then be used to compare the viremia and brain infectivity levels of an attenuated WNV isolate in comparison to a highly neuroinvasive WNV strain, WN-NY99. Viral titers in both serum and brain, as measured during a time course experiment, will demonstrate the hypothesized variation in viremia induced by each virus.

Specific aim 3

Identification of nucleotides/amino acid substitutions conferring phenotypic variation in North American WNV by reverse genetics.

Hypothesis: *Specific mutations to the WNV genome are responsible for the phenotypic variation observed in 2003 WNV isolates.*

Rationale: RT-PCR has been used to sequence the complete genome of one of the 2003 attenuated WNV isolates revealing several unique mutations in this isolate. Sequence analysis of specific genes/regions other 2003 attenuated WNV isolates will identify those mutations shared by attenuated isolates. Following the identification of potential

nucleotide/amino acid substitutions that correlate with the phenotypic variation observed in 2003 WNV isolates, site-directed mutagenesis of an infectious clone based on WN-NY99 (382-99) will be performed in order to map the genetic determinants of changes in the WNV phenotype. Based on preliminary sequencing of the prM and E protein genes, it is expected that the mutations conferring the phenotypic variation reside in nonstructural protein genes.

Figure 1-1. Worldwide distribution of West Nile virus, 2005. Kunjin virus in Australia appears in light blue. (Modified from CDC, Arbonet).

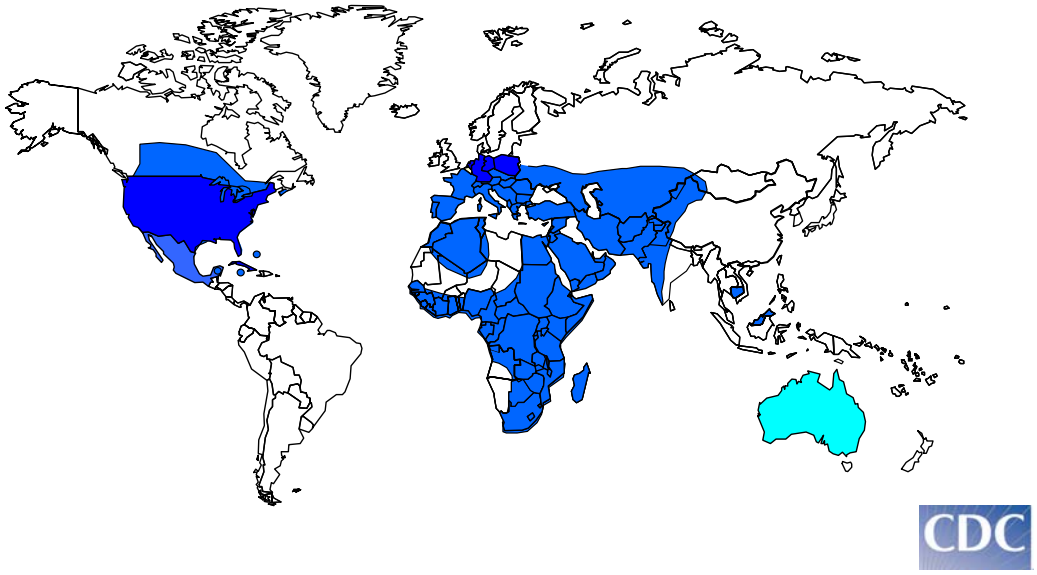


Figure 1-2. Unrooted cladogram showing relationships of Old World WNV isolates by lineage (Modified from Berthet et al., 1997).

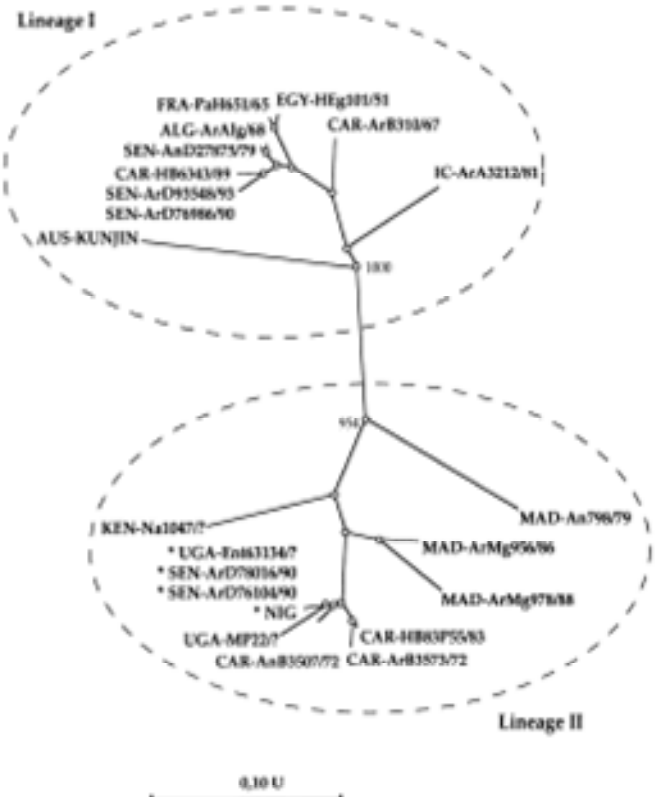


Figure 1-3. Schematic showing the major components in the transmission and maintenance of arboviral encephalitis (adapted from CDC, Arbonet).

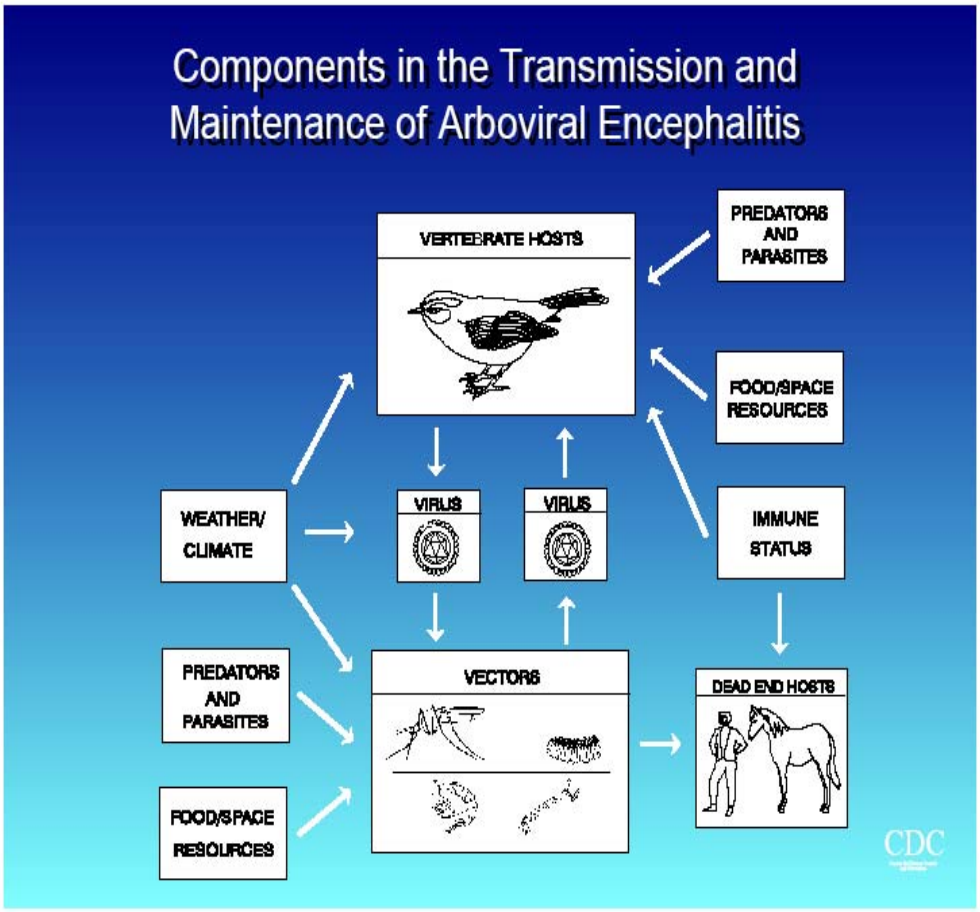
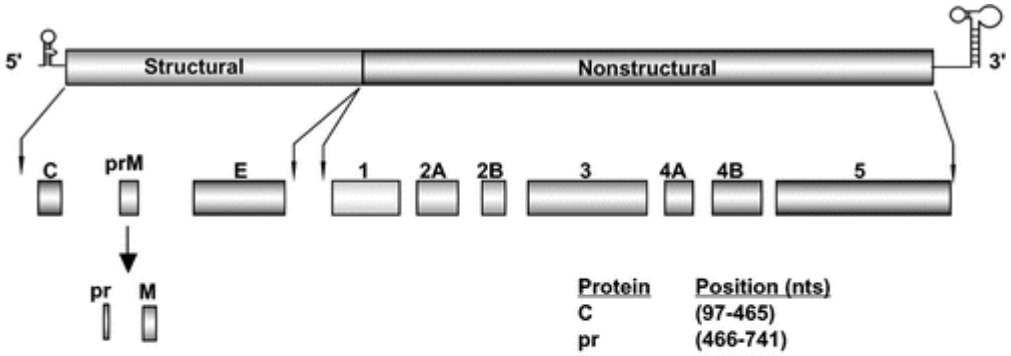


Figure 1-4. West Nile virus genome organization (modified from Brinton, 2002).



Protein	Position (nts)
C	(97-465)
pr	(466-741)
M	(742-966)
E	(967-2469)
NS1	(2470-3525)
NS2A	(3526-4218)
NS2B	(4219-4611)
NS3	(4612-6468)
NS4A	(6469-6915)
NS4B	(6916-7680)
NS5	(7681-10395)

Figure 1-5. Ribbon diagram of the DEN E protein. A: Domain definition showing domains I, II, and III in red, yellow, and blue, respectively. B: Top view of E protein homodimer (modified from Modis et al., 2003).

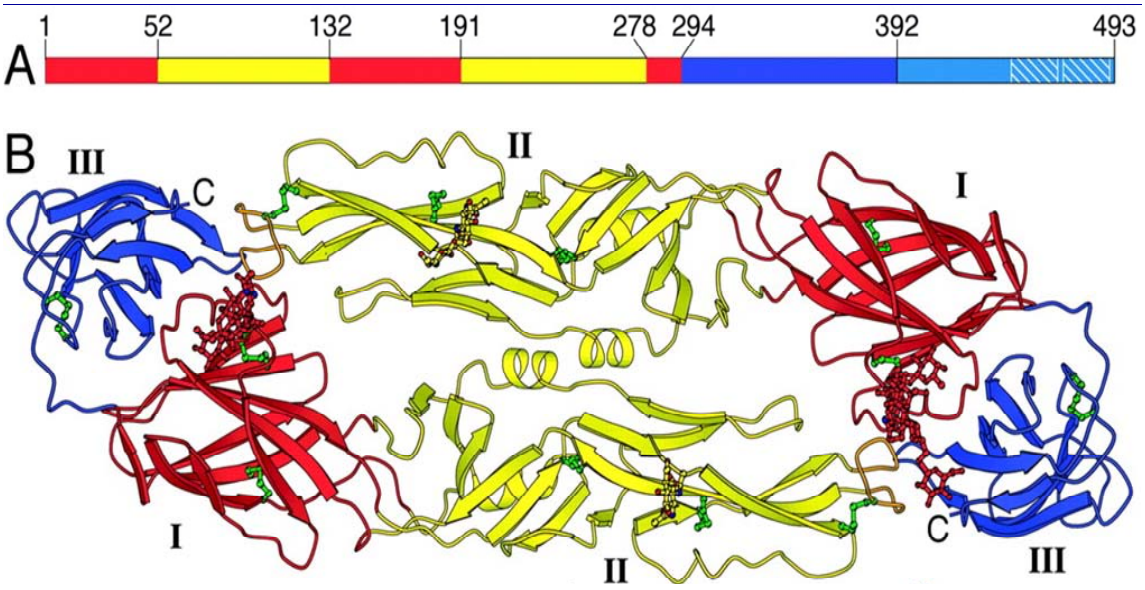


Figure 1-6. West Nile virus life cycle (Modified from Mukhopadhyay et al., 2005).

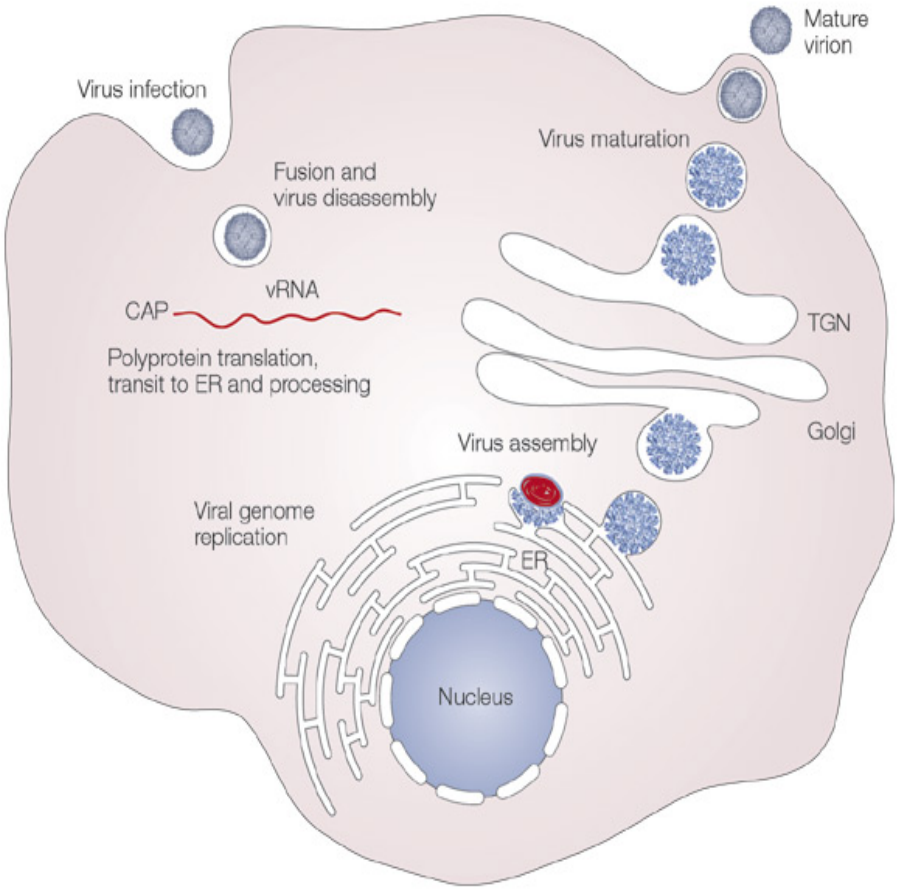


Figure 1-7. Reported incidence of neuroinvasive West Nile virus disease by county, United States, 1999–2004. Reported to Centers for Disease Control and Prevention by states through April 21, 2005 (Adapted from Hayes et al., 2005).

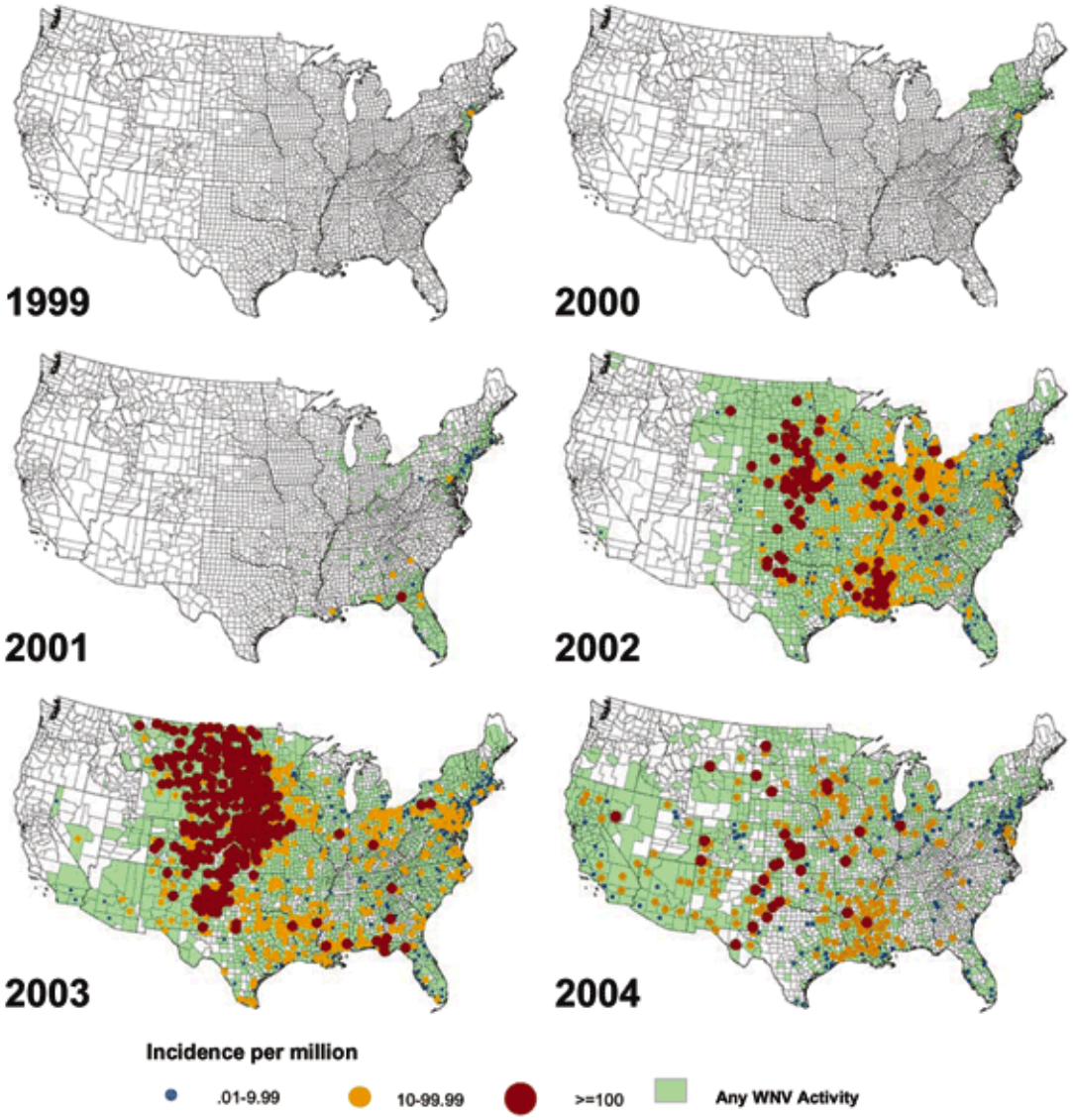


Figure 1-8. Distribution of human cases of West Nile virus in the U.S. by year (CDC, Arbonet).

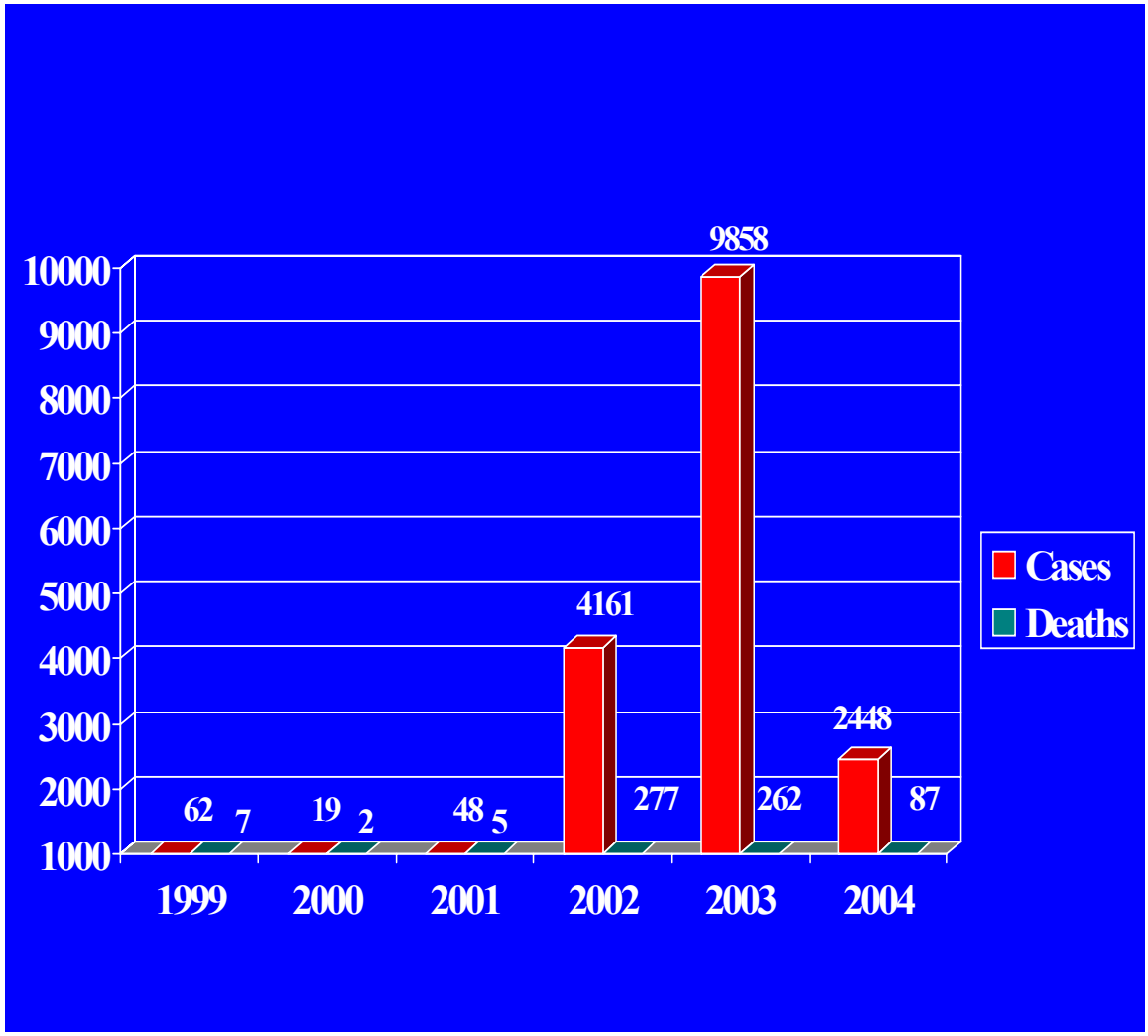


Figure 1-9. Distribution of equine cases of West Nile virus in the U.S. by year (CDC, Arbonet).

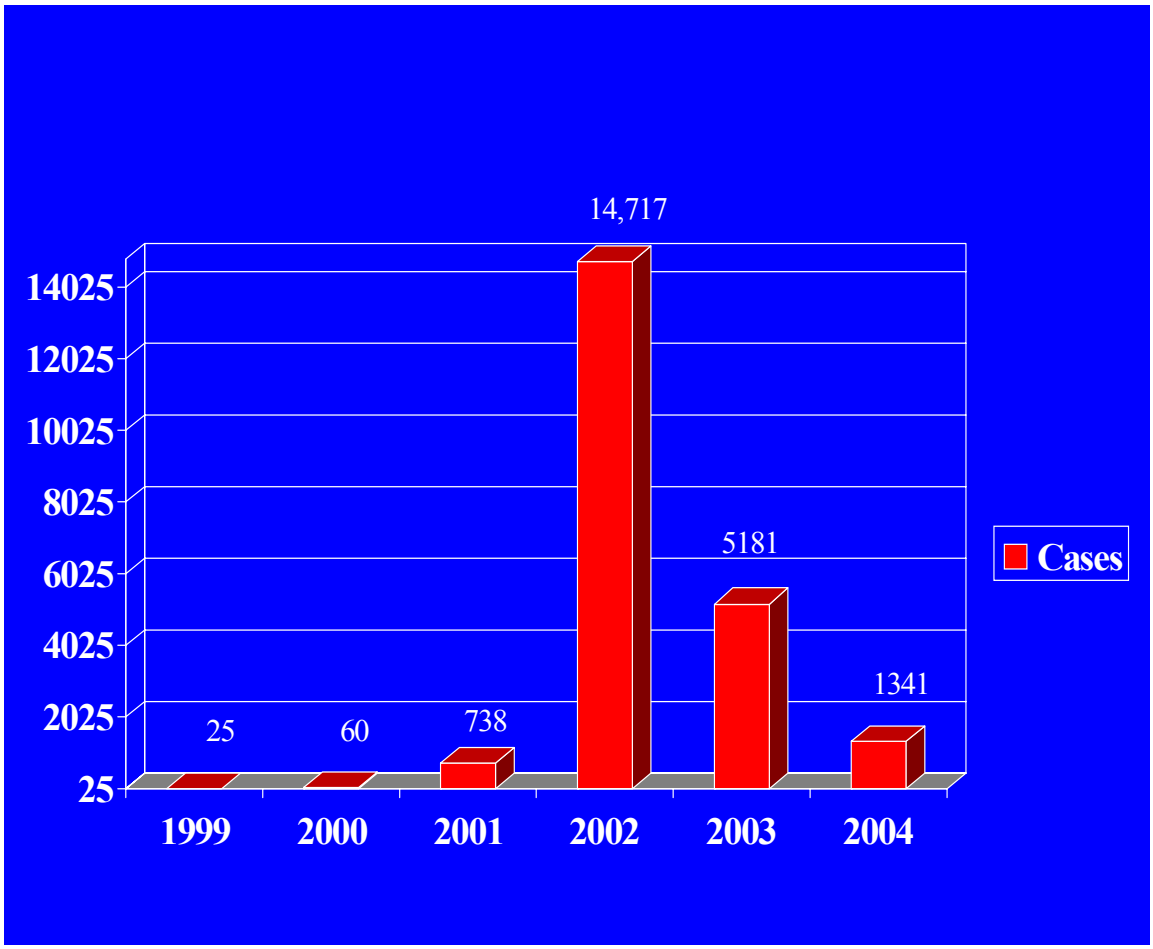
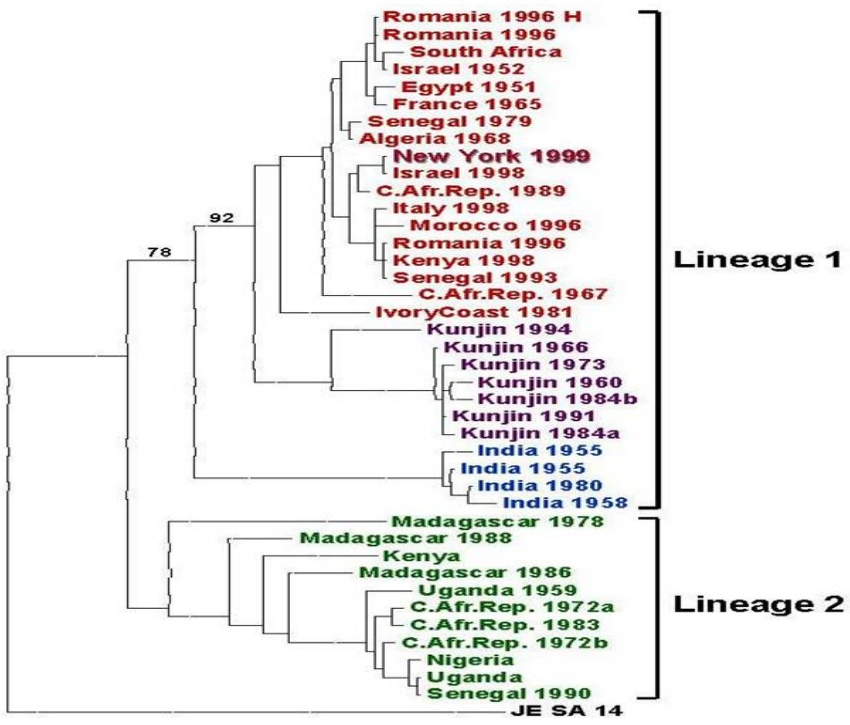


Figure 1-10. Genetic lineages of West Nile virus based on E protein gene alignment. Colors represent individual subclades (Adapted from Lanciotti et al., 1999).



CHAPTER 2

MATERIAL AND METHODS

Virus isolation and passage

Isolates were collected from five states including Illinois, Alabama, Louisiana, Colorado, and Texas. Isolates from Texas were collected from nine counties representing regions across the entire state (Fig. 2-1, 2-2). All isolates were collected between September 2001 and October 2002. After being confirmed WNV positive by each of the above agencies, virus or tissues were sent to UTMB for submission into the World Reference Center for Emerging Viruses and Arboviruses. Dead birds and mosquitoes were collected by the Harris County Mosquito Control Division from Harris and Montgomery Counties, TX during the summer of 2003. Virus isolations were made by inoculation of bird brain or mosquito homogenates in Vero cells as described elsewhere (Lillibridge et al., 2004). Upon confirmation that isolates were WNV positive by either hemagluttination inhibition assay or RT-PCR, each isolate was passaged a single time in Vero cells to derive virus for use in these studies and was submitted to the World Reference Center for Emerging Viruses and Arboviruses at the University of Texas Medical Branch. Virus samples represented a variety of sources including mosquito pools, bird brain, human cerebrospinal fluid (CSF), and a dog kidney. Of the 18 isolates

collected in 2001 and 2002 (Table 2-1), eleven were isolated from mosquito pools by the Texas Department of Health (TDH); two from a mosquito pool and dog kidney homogenate by the Illinois Health Service (INHS); two from passerine brain homogenates from the University of Alabama at Birmingham; one from a red-tailed hawk brain homogenate by the Centers for Disease Control and Prevention-Division of Vector-Borne Infectious Disease (CDC-DVBID), Colorado; one from a mosquito pool in Louisiana courtesy of the CDC-DVBID; and one from the CSF of a patient that died of West Nile encephalitis at UTMB. All newly sequenced WNV isolates came from a variety of sources including birds, mosquitoes, horses, and humans. A total of 74 isolates from 2003 and 2004 were additionally collected from locations including Texas, Nebraska, Ohio, Indiana, Illinois, Georgia, Florida, Louisiana, Arizona, California, Colorado, Mexico and Canada (Figure 2-2; Table 2-2).

Viral RNA extraction

Viral RNA was extracted directly from 140 μL of infected Vero or BHK cell culture supernatants using the QiaAMP viral RNA extraction kit according to protocol provided by manufacturer (Qiagen).

Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

RT was performed in a 50 μL volume containing 5 μL of viral RNA, 1 μL of random hexamer primer, 10 μL of 5X RT buffer, 4 μL of 10 mM dNTPs, 0.4 μL of cloned RNAse inhibitor, 0.5 μL of Moloney murine leukemia virus (MMLV) reverse transcriptase, and 29.1 μL of HPLC purified water. RNA and primer were first heated at 72°C for 10 minutes, followed by addition of reaction reagents for 10 minutes at 55ºC and 60 minutes at 45ºC. RNA was stored at -80º C until further use. PCR was performed in a 25 μL volume containing 2.0 μL cDNA template from RT, 1.0 μL forward primer, 1.0 μL reverse primer, 2.5 μL 10X PCR buffer, 0.5 μL 10 mM dNTPs, 0.5 μL of 1 U/μL Taq polymerase, and 17.5 μL of HPLC water. Three primer pairs were used to amplify the entire prM-E genes of WNV (Table 2-3). PCR conditions were as follows:

94ºC for 5 minutes – 1x

94ºC for 40 seconds
54ºC* for 40 seconds
72ºC for 3 minutes
} 34x

72ºC for 7 minutes – 1x

* Annealing temperatures varied depending on the melting temperatures of each primer used for the reaction.

Each reaction was cooled to 4ºC and stored until PCR products were separated by gel electrophoresis in 2% agarose gels containing ethidium bromide (0.5 µl/ml in TAE buffer). Following visualization under UV light, PCR bands were cut out of gels and purified for sequencing/cloning using a Gel Extraction Kit according to the manufacturer's instructions (Qiagen).

Primers and primer design

Primers were designed based on the prototypical North American West Nile virus strain, NY99 (382-99), which was isolated from the brain of a dead flamingo in the Bronx Zoo in 1999. The complete genome of this strain was previosly sequenced by Lanciotti et al. (1999) and stored in GenBank (accession no. AY196835). A set of eight primers with varying annealing temperatures were made to allow a consensus sequence to be constructed from overlapping amplicons aligned using the ContigExpress program in the Vector NTI software package (Informax, Version 8.0). RT-PCR primers were used as sequencing primers. A complete list of primers used is found in Table 2-3.

Nucleotide sequencing/phylogenetic analysis of 2001-2002 isolates

prM and E genes

Following gel purification of PCR products, the resulting template was directly sequenced using the amplifying primers. Sequencing reactions were performed in the UTMB Biomolecular Resource Facility's DNA sequencing laboratory using the ABI

PRISM Big Dye Terminator v3.0 cycle sequencing kits (Applied Biosystems) according to the manufacturer's protocol and analyzed on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Analysis and assembly of sequencing data were performed using the Vector NTI Suite software package (Informax). A total of 22 isolates from 2001 and 2002 were sequenced by the author. Nucleotide and deduced amino acid sequences of the entire prM-E genes from each isolate were aligned using the AlignX program in the Vector NTI Suite and compared to previously published sequences of isolates from southeast Texas collected between June and August of 2002 (Beasley et al., 2003). All isolates were then compared to isolates collected in the northeastern U.S. during 1999, 2000, and 2001, and a phylogenetic tree was constructed by maximum parsimony algorithm using PAUP (Version 4.0b10) to show genetic relationships of these isolates to other North American WNV isolates found in GenBank in which the homologous 2004 nucleotide region had been sequenced.

Nucleotide sequencing/phylogenetic analysis of 2003-2004 isolates and complete genome sequencing of phenotypic variants

prM and E, individual genes, and complete genome sequencing

The prM and E genes of 43 isolates from 2003 and 2004 and the complete genomes of 5 isolates (2003 and 2004) were sequenced by the author using RT-PCR as described above from RNA extracted from infected tissues, cell culture supernatants of original isolations, or supernatants after a single passage in African green monkey (Vero) cell culture. Additional isolates used in the phylogenetic analysis were generously

provided by Brault, Ebel, and Lanciotti (personal communication). The Titan single-step RT-PCR kit (Roche) was also used for DNA amplification according to the manufacturer's protocol to expedite complete genome sequencing. Briefly, each reaction mixture containing RNA, forward and reverse primers, and PCR reagents, as described above, were heated at 50°C for 30 minutes for reverse transcription, directly followed by PCR. Reaction conditions were as follows:

94ºC for 2 minutes – 1x

94ºC for 10 seconds
54ºC* for 30 seconds
68ºC for 1 minute
} 10x

94ºC for 10 seconds
54ºC* for 30 seconds
68ºC for 45 seconds
} 25x
+ cycle elongation of 5 seconds for each cycle

68ºC for 7 minutes – 1x

* Annealing temperatures varied depending on the melting temperatures of each primer used for the reaction.

Primers used for all reactions are listed in Table 2-3. Following completion of either partial or complete genome sequencing, nucleotide and deduced amino acid sequences were aligned with sequences from other WNV isolates found in GenBank that

represented homologous sequence regions. The AlignX program of the Vector NTI Suite software package (Informax) was used to generate all alignments. The GenBank accession numbers for isolates used in the generation of phylogenetic trees are found in each table describing source of isolates (Table 2-2). Bayesian analysis was used to generate a phylogenetic tree based on the prM-E alignment of 74 North American WNV isolates sequenced by the author and others with other North American WN viruses available in GenBank and rooted with the most closely related Old World WNV, Israel-1998 (Accession no. AY033389), in order to create a tree illustrating more parsimony informative sites. Phylogenies were generated by the program MRBAYES (Version 2.0) (Huelsenbeck and Ronquist, 2001) using the Metropolis coupled Markov chain Monte Carlo algorithm run with four chains over 150,000 generations under a general time-reversible model with a burn-in time of 50,000 generations. Rate heterogeneity was estimated using a γ distribution for the variable sites. The consensus phylogram among the 108 WNV isolates analyzed was then generated from the MRBAYES output file using PAUP (Version 4.0b11, Sinauer Associates) with clade credibility values at relevant nodes to demonstrate statistical support for each clade. The Bayesian consensus tree was compared to trees generated by neighbor-joining and maximum parsimony analyses using PAUP in order to determine if similar tree topologies were generated using different methods. Additional phylogenetic trees were generated from alignments of complete genomes or individual genes/untranslated regions (5'UTR, Capsid, prM, E, NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5, and 3'UTR) comparing the 5 genomic sequences determined by the author to all other WNV complete genome sequences

available in GenBank and from collaborators. Maximum likelihood analysis was used to generate all additional trees using PAUP under the general time-reversible model with a γ distribution of substitution rates. Confirmation of tree topology was done using maximum parsimony and neighbor-joining methods with statistical support for relevant clades provided by 500 bootstrap replicates using the neighbor-joining analysis.

Cloning and quasispecies analysis

The WN1751/WN2504A PCR product derived from WNV isolate Galveston Co., TX-3 was cloned into pGEM-T Easy (Promega Corporation) and ten clones were sequenced to determine the degree of nucleotide sequence divergence within a single isolate collected from the southeast coast of Texas.

Cell culture techniques

Each isolate made for sequencing studies were given a single passage in Vero cells to derive at least 12 x 0.5ml ampoules of virus for use in these studies. Virus ampoules were stored at -80.0 °C until thawed for experiments. Each ampoule was used only once in order to insure accuracy of virus titer.

Viral plaque morphology assays and titration

In order to establish the plaque morphology and titer of isolates collected in 2003 and 2004, plaque assays were carried out by infecting Vero cells in 6-well plates with

WNV isolates at dilutions of 10^{-1} to 10^{-6}. All plaque assays were run in parallel with WN-NY99 as a positive control and PBS as a negative control. Plaque morphology was measured by staining plaque assays with crystal violet followed by digital photography of plaques. Images were visualized in Microsoft Photoshop and diameters of plaques were measured from these images. The small plaque (sp) phenotype is described as a < 1.0 mm difference in plaque diameter compared to the control WN-NY99 large plaque (lp) size > 1.5mm.

Temperature-sensitivity assays

Each isolate was also plaqued at 37.0 °C, 39.5°C, and 41.0 °C for 72 hours to measure temperature-sensitivity (ts). The ts phenotype is described as a >2.0 $\log_{10}$ PFU reduction in titer determined by Vero cell plaque assay at either 39.5°C or 41.0°C when compared to permissive temperature 37.0 °C.

In vitro replication kinetic assays

Growth curves were performed in triplicate by infecting Vero cells in 12-well plates with WNV isolates at an MOI of 0.1 PFU/cell. A low MOI was used in order to compare the multiplication kinetics of the variant viruses as a contributing factor to their attenuated phenotype rather than to compare their replication kinetics in a single step growth curve. Supernatants were then harvested at 0, 12, 24, 48, 72, and 96 hours post-

infection. Vero cell plaque assays were used to determine mean virus titers at each time point.

Mouse neuroinvasiveness/neurovirulence screening

Screening WNV isolates for attenuation of neuroinvasiveness was performed by intraperitoneal (ip) inoculation of groups of five female 3-4 week-old Swiss Webster mice with 100 µl of 10^2 and 10^3 pfu of virus. WN-NY99 (isolate 385-99) was used as a positive control; mice inoculated with saline served as negative controls. Mice were evaluated daily for signs of illness (ruffled fur, hunched posture, lethargy, ocular or nasal discharge, hindlimb/forelimb paralysis) or death. All deaths occurred between twice-daily observations. At 21 days post-inoculation (dpi) surviving mice were challenged with 100 LD_{50} of WN-NY99 (385-99) to ensure that replication of virus had occurred in each mouse inoculated, as indicated by protection following challenge with a lethal dose. In order to determine if isolates that were attenuated in neuroinvasiveness were also attenuated for neurovirulence, attenuated isolates (20 µl) were inoculated intracerebrally into groups of five female 3-4 week-old Swiss Webster mice at dosages of 10^3 to 10^{-1} pfu of virus. Severe illness and/or death were the end-points as before. All procedures with animals were carried out according to guidelines of the Committee on Care and Use of Laboratory Animals under an animal-care protocol approved by the University of Texas Medical Branch. All work with virus-infected animals and was carried out in Biosafety Level 3 facilities.

LD_{50} determination

In order to establish the extent of attenuation of neuroinvasiveness in a Swiss Webster mouse model, a precise LD_{50} for WNV isolates that exhibited a ≥1,000 pfu LD_{50} was determined. This was done by intraperitoneal (ip) inoculation of groups of five female 3-4 week-old Swiss Webster mice with 100 µl of serial ten-fold dilutions of virus from 10^{6} to 10^{-1} pfu of virus. WN-NY99 (385-99) was used as a positive control. Mice were evaluated for signs of illness or death, as before. At 21 days post-inoculation (dpi), surviving mice were challenged with 100 LD_{50} of WN-NY99 (385-99) to ensure replication of virus and to calculate the dose that protects 50% of mice from death against a lethal challenge of WN-NY99 (PD_{50}).

Serum viremia/brain infectivity studies

The 3-4 week-old Swiss Webster mouse model was used to compare the viremia and brain infectivity levels of an attenuated WNV isolate to the highly neuroinvasive WNV isolate, WN-NY99. Two groups of 24 mice were infected ip with 100 µl of 10^{3} pfu of either WN-NY99 (385-99) or an attenuated 2003 WNV isolate (Bird 1153) and monitored daily for signs of illness. Following infection, three mice from each of the two groups were bled daily for eight days and serum samples were titrated, using standard Vero cell viral plaque assays in six well plates. Briefly, animals were bled daily by cardiac puncture after euthanasia in order to recover at least 2 ml of whole blood. Blood

was allowed to clot and then centrifuged; the serum was extracted and diluted 10^{-1} to10^{-6} for titration by plaque assay. Virus titers in brain were measured daily by removing the entire brain from each of three euthanized mice, followed by homogenization in 2% MEM. Brain homogenates were then diluted 10^{-1} to10^{-6} for titration by plaque assay. Virus titers in blood and brain at each time interval were measured in three mice and the averages of these were used to construct a viremia and brain infectivity curve. These results were compared to a viremia curve and brain infectivity curve for WN-NY99, which was performed in parallel under the same experimental conditions. All procedures were carried out according to guidelines of the Committee on Care and Use of Laboratory Animals under an animal-care protocol approved by the University of Texas Medical Branch. All work with virus-infected animals was carried out in Biosafety Level 3 facilities.

West Nile virus infectious clone technology

The WNV NY99 virus-specific infectious cDNA clone was constructed in two plasmids, utilizing a derivative of plasmid pBRUC-139S (Figure 2-4). Plasmid pWN-AB contained WNV nucleotides 1 to 2495, which were preceded by restriction sites SstI and MluI and the promoter for T7 polymerase. Plasmid pWN-CG contained WNV nucleotides 2495 to 11029 and an engineered 3'-terminal XbaI site for plasmid linearization just prior to transcription of genomic RNA. Nucleotide sequencing of the NY99 infectious clone-derived virus (NY99ic) and the parental NY99 stock from which

it was derived identified seven nucleotide differences from the published NY99 sequence (GenBank accession no. AF196835), including two that encoded amino acid substitutions: C1428U, U1855C, C3880U (NS2A-118 His to Tyr), A4922G (NS3-104 Lys→Arg), G7029U, U8811C, and A10851G. Full-genomic-length cDNA was prepared by cleaving the pWN-AB and pWN-CG plasmids at the natural NgoMIV- nucleotide 2495 site of WNV followed by ligating the two plasmids at this NgoMIV site. The in vitro-ligated DNA fragment containing the full-genome-length WNV cDNA was then purified by phenol/chloroform extraction and pelleted following an overnight ethanol precipitation. Following resuspension of the purified cDNA in TE buffer (pH 8.0), viral genomic RNA was transcribed by using the AmpliScribe T7 kit (Epicentre Technologies). Transcription was carried out in the presence of m^7-GpppA cap analog for 2 to 3 h at 37°C, and Vero cells were transfected with the transcribed RNA by electroporation in 0.2 cm electrode gap cuvettes (Biorad) at 1.5kV, infinite Ohms, and 25μF. Transfections were then transferred to T_{75} flasks with 8% MEM and observed daily for CPE. Rescued virus was harvested following the appearance of CPE or after 6 days post-infection.

Site-directed mutagenesis using the infectious clone

In order to identify mutations to the WNV genome conferring phenotypic variation, site-directed mutagenesis of the infectious clone was used to substitute an amino acid that was identified by sequencing studies to differ between WN-NY99 and isolates displaying phenotypic variation. Results from Chapter 5 have identified several

isolates with sp, ts, and attenuated phenotypes that also share amino acid substitutions when compared to WN-NY99. The substitution at NS4B (E249G), which is shared by three of the sp, ts, and attenuated 2003 WNV isolates, was the starting point at which site-directed mutations were made to the infectious clone. Mutations to the infectious clone were made using the QuikChange® XL Site-Directed Mutagenesis Kit (Stratagene). Briefly, this system can be used to make amino acid substitutions in the infectious clone by designing mutagenic primers which anneal to the region of interest of the WNV DNA inserted in the plasmid (e.g., nucleotide 7666 [NS4B E249G] of the WNV genome). The plasmid from the infectious clone containing the insert is denatured, allowing the mutagenic primers to anneal to the region of interest. The primers, which are each complementary to opposite strands of the plasmid insert, are extended during temperature cycling using *PfuTurbo* DNA polymerase. Incorporation of the mutagenic primers produces a mutated plasmid containing staggered nicks. When the temperature cycling is complete, the product is treated with *Dpn* I (an endonuclease which recognizes and digests methylated and hemimethylated DNA). Digestion with *Dpn* I digests the parental DNA template and selects for only synthesized DNA containing the mutation of interest. The synthesized DNA is then transformed into XL10-Gold ultracompetent cells to allow for nicked ends of the plasmid DNA to be repaired and to produce the mutated plasmid in larger quantities. Following generation of the desired plasmid, the reverse genetics system described above was used to generate an infectious WNV with the same mutation found in the sp, ts, and attenuated WNV isolates. Because it was possible that the observed phenotypes of 2003 WNV isolates were the result of an accumulation of

mutations in the genome, it was necessary to continue to add mutations to the infectious clone in order to generate virus with measurable phenotypic variation. Mutations to the infectious clone (those unique to sp, ts, and attenuated isolates) were made as both single amino acid substitutions and as combinations of substitutions in order to account for the possibility of accumulated mutations resulting in specific viral phenotypes.

3'UTR exchange between isolate Bird 1153 and WN-NY99 infectious clone

In order to completely exchange the 3'UTR of WNV strain Bird 1153 with the 3'UTR of the NY99 infectious clone, the 3'UTR of Bird 1153 was amplified by reverse transcriptase PCR (RT-PCR) using primers that introduced a 5' SalI site (3'UTR For: CAACTTTGGTCGACGACACAGTACTGT) and a 3' XbaI site (3'UTR Rev: TCTAGAAGATCCTGTGTTCTCGCACC). This PCR product was then cloned into pGEM-T for substitution into the pWN-CG plasmid of the NY99 infectious clone. In order to do an exact exchange of the Bird 1153 3'UTR into the NY99 backbone, the Quikchange site-directed mutagenesis kit (Stratagene) was used using the same primers to engineer those restriction sites at either end of the pWN-CG 3'UTR. None of these mutations resulted in changes to the nucleotide sequence as the restriction site motifs in the 3'UTR forward and reverse primers were contained outside of the WNV nucleotide sequence.

Infectious clone virus rescue

In order to confirm that infectious virus was recovered following transfection, cell cultures were followed for a period ranging from 3-6 days post-transfection and observed for the presence of CPE. Upon indication of sufficient CPE, RT-PCR was performed on cell culture supernatants using primers that amplified the prM and E protein genes of the WNV genome. If the RT-PCR produced a positive result, the cell culture supernatants were harvested into at least 12 x 0.5ml ampoules and stored at -80.0 °C for future use.

Genetic characterization of mutant viruses

For those transfections that produced positive RT-PCR results, primers specific to regions containing the desired mutation(s) were used to amplify and sequence the PCR product to determine if the desired mutation was incorporated into the genome of the virus.

Phenotypic characterization of mutant viruses

Upon successful viral rescue and and confirmation of site-directed mutations to the viral genome, a plaque morphology assay and ts assay were used to examine the phenotypic characteristic of each isolate. Viral multiplication kinetic studies were undertaken to measure the multiplication capacity of mutant viruses displaying sp and/or ts phenotypes. The ip LD50 of each mutant virus was also measured using the same 3-4 week old female Swiss Webster mouse model as described above.

Table 2-1. List of isolates used for microevolution studies (Chapter 3).

Strain	Source	Collected	RNA origin
WN-NY99	Flamingo	1999	Brain
(AF196835)			
Harris Co., TX	Bluejay	06/11/02	Brain/Vero
(AY185906)			
Harris Co., TX	Bluejay	06/10/02	Brain/Vero
(AY185907)			
Nueces Co, TX - 1	*C. quinquefasciatus*	08/06/02	BHK
Nueces Co., TX - 2	*C. quinquefasciatus*	09/17/02	BHK
Gregg Co., TX	*C. quinquefasciatus*	09/25/02	BHK
Tarrant Co., TX	*C. restuans*	09/30/02	Vero
Wichita Co., TX	*C. quinquefasciatus*	10/23/02	BHK
Randall Co., TX	*C. tarsalis*	09/18/02	BHK
El Paso Co., TX	*C. tarsalis*	08/26/02	Vero
Illinois - 1	*C. pipiens*	08/02/02	Vero
Illinois - 2	Dog	08/01/02	Kidney/Vero
Alabama - 1	*C. quinquefasciatus*	10/05/01	Vero
Alabama -2	Crow	09/10/01	Brain/Vero
Colorado	Red-tailed Hawk	08/01/02	Brain/Vero
Louisiana	*C. salinarius*	08/06/02	Vero
Galveston Co., TX-1	Bluejay	08/02/02	Vero
(AY185914)			
Galveston Co., TX-2	Bluejay	07/19/02	Vero
(AY185913)			
Galveston Co., TX-3	*C. quinquefasciatus*	08/21/02	Vero
Jefferson Co., TX - 1	*C. quinquefasciatus*	08/06/02	BHK
Jefferson Co., TX - 2	*C. quinquefasciatus*	07/02/02	BHK
Jefferson Co., TX - 3	Human	08/24/02	CSF/Vero
Orange Co., TX	*C. quinquefasciatus*	07/03/02	Vero

Table 2-2. Isolates used for North American phylogenetic analysis (Chapter 4).

Abbreviation	Isolate	Year of isolation	Location	Source	Passage History	Genbank Accession No.
FL 2001 Palm Beach	Florida 2001	2001	Palm Beach, Florida	Catbird	V2	DQ080072
NY 2001 Suffolk	32010157	2001	Suffolk Co., NY	*Cx. pipiens/restuans*	V1	DQ164194
FL 2002 Sumter	Florida 2002	2002	Sumter, Florida	Horse	V2	DQ080071
NY 2002 Queens	02003011	2002	Queens Co., NY	American crow	P	DQ164186
NY 2002 Broome	02003557	2002	Broome Co., NY	American crow	P	DQ164187
NY 2002 Clinton	02002758	2002	Clinton Co., NY	American crow	P	DQ164193
NY 2002 Nassau	34020055	2002	Nassau Co., NY	*Cx. pipiens/restuans*	V2	DQ164195
Ontario 2002	Ontario	2002	Ontario, Canada	American crow	V2	DQ158245
Saskatchewan 2002	Saskatchewan	2002	Saskatchewan, Canada	Black-billed magpie	V2	DQ158246
Manitoba 2002	Manitoba	2002	Manitoba, Canada	Amercican crow	V2	DQ158244
GA 2002 1	68955	2002	Georgia	Human - plasma	P	DQ164196
GA 2002 2	68960	2002	Georgia	Human - brain	P	DQ164197
OH 2002	81948	2002	Ohio	Human - plasma	P	DQ164202
IN 2002	81931	2002	Indiana	Human - plasma	P	DQ164200
TX 2002 1	80025	2002	Texas	Human - plasma	P	DQ164198
TX 2002 2	80022	2002	Texas	Human - plasma	P	DQ164205
NY 2003 Westchester	03000360	2003	Westchester Co., NY	Amercican crow	P	DQ164188
NY 2003 Alabany	03001986	2003	Albany Co., NY	American crow	P	DQ164189
NY 2003 Suffolk	03002018	2003	Suffolk Co., NY	American crow	P	DQ164190
NY 2003 Chautauqua	03002086	2003	Chautauqua Co., NY	American crow	P	DQ164191
NY 2003 Rockland	03002094	2003	Rockland Co., NY	American crow	P	DQ164192
AZ 2003 1	03-az-mp-1623	2003	Arizona	Magpie	V1	DQ158207
AZ 2003 2	03-az-mp-1681	2003	Arizona	Magpie	V1	DQ158208
AZ 2003 3	03-az-mp-1799	2003	Arizona	Magpie	V1	DQ158209
CA 2003 Los Angeles 1	03-ca-crow-s0331532	2003	Los Angeles Co., CA	American crow	V1	DQ158210
CA 2003 Los Angeles 2	03-ca-crow-s0334814	2003	Los Angeles Co., CA	American crow	V1	DQ158211
CA 2003 Los Angeles 3	Mosq. grla1131	2003	Los Angeles Co., CA	*Culex tarsalis*	V2	DQ158212
CA 2003 Los Angeles 4	03-ca-mp-grla-1260	2003	Los Angeles Co., CA	Magpie	V1	DQ158214
CA 2003 Los Angeles 5	Bird 9173	2003	Los Angeles Co., CA	American crow	V2	DQ158217
CA 2003 San Bernadino	Bird 9172	2003	San Bernadino Co., CA	American crow	V2	DQ158218
CA 2003 1	03-ca-mp-impr-102	2003	Imperial Valley, CA	Magpie	V1	DQ158215
CA 2003 2	03-ca-mp-impr-1075	2003	Imperial Valley, CA	Magpie	V1	DQ158216
CA 2003 3	Mosq. impr1143	2003	Imperial Valley, CA	*Culex tarsalis*	V2	DQ158213
MX 2003 Nuevo Leon	Mosq. 9488	2003	Nuevo Leon, Mexico	*Culex quinquefasciatus*	V2	AY963775

Table 2-2 (continued). Isolates used for North American phylogenetic analysis (Chapter 4).

Abbreviation	Isolate	Year of isolation	Location	Source	Passage History	Genbank Accession No.
CO 2003 Larimer	Bird 9185	2003	Larimer Co., CO	*Culex pipiens*	V2	DQ158219
CO 2003 1	Colorado 3068	2003	Colorado	Red-tailed hawk	P	DQ164204
CO 2003 2	Colorado 3258	2003	Colorado	Magpie	P	DQ164203
NB 2003 Brown	Bird 9239	2003	Brown Co., Nebraska	Bluejay	V2	DQ158236
NB 2003 Dakota	Bird 9241	2003	Dakota Co., Nebraska	Bluejay	V2	DQ158237
TX 2003 Harris 1	Bird 9045	2003	Harris Co., TX	Bluejay	V1	DQ158223
TX 2003 Harris 2	Bird 9114	2003	Harris Co., TX	Bluejay	V1	DQ158222
TX 2003 Harris 3	Mosq. V4095	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	DQ158224
TX 2003 Harris 4	Mosq. V4369	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	AY712948
TX 2003 Harris 5	Mosq. V4096	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	DQ158226
TX 2003 Harris 6	Bird 1153	2003	Harris Co., TX	Mourning dove	V1	AY712945
TX 2003 Harris 7	Bird 1171	2003	Harris Co., TX	Great-tailed grackle	V1	AY712946
TX 2003 Harris 8	Bird 1175	2003	Harris Co., TX	Bluejay	V1	DQ158220
TX 2003 Harris 9	Bird 1240	2003	Harris Co., TX	Bluejay	V1	DQ158221
TX 2003 Harris 10	Bird 1461	2003	Harris Co., TX	Bluejay	V1	AY712947
TX 2003 Harris 11	Mosq. V4370	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	DQ158225
TX 2003 Montgomery 1	Bird 1519	2003	Montgomery Co., TX	Bluejay	V1	DQ158227
TX 2003 Montgomery 2	Bird 1574	2003	Montgomery Co., TX	Bluejay	V1	DQ158228
TX 2003 Montgomery 3	Bird 1576	2003	Montgomery Co., TX	Bluejay	V1	DQ158229
TX 2003 Jefferson 1	Bird 1881	2003	Jefferson Co., TX	Mourning Dove	V1	DQ158230
TX 2003 Jefferson 2	Bird 2073	2003	Jefferson Co., TX	Bluejay	V1	DQ158231
TX 2003 Wharton	Bird 2071	2003	Wharton Co., TX	Bluejay	V1	DQ158232
TX 2003	Texas 82229	2003	Texas	Human - plasma	P	DQ164199
LA 2004 New Iberia	Bird 2409	2004	New Iberia Parish, Louisiana	Northern cardinal	V2	DQ080061
IL 2004 Madison	Illinois - 9515	2004	Madison Co., Illinois	American crow	V2	DQ158250
IL 2004 Schuyler	Illinois - 9517	2004	Schuyler Co., Illinois	American crow	V2	DQ158249
IL 2004 Knox	Illinois - 9519	2004	Knox Co., Illinois	Bluejay	V2	DQ158248
IL 2004 New Marion	Illinois - 9520	2004	New Marion Co., Illinois	Bluejay	V2	DQ158247
AZ 2004	Arizona 2004	2004	Arizona	Human-plasma	P	DQ164201

Table 2-2 (continued). Isolates used for North American phylogenetic analysis (Chapter 4).

Abbreviation	Isolate	Year of isolation	Location	Source	Passage History	Genbank Accession No.
MX 2004 Sonora	Mexico 2004	2004	Sonora, Mexico	Human-plasma	V2	AY963774
TX 2004 Harris 1	Bird 2419	2004	Harris Co., TX	Bluejay	V1	DQ158233
TX 2004 Harris 2	Bird 2541	2004	Harris Co., TX	Mourning dove	V1	DQ158234
TX 2004 Harris 3	Bird 3218	2004	Harris Co., TX	Bluejay	V1	DQ158235
TX 2004 Harris 4	Bird 3588	2004	Harris Co., TX	Bluejay	V1	DQ164206
TX 2004 Parmer	Bird 9460	2004	Parmer Co., TX	House sparrow	V1	DQ158238
TX 2004 Brazoria 1	Bird 9461	2004	Brazoria Co., TX	Bluejay	V1	DQ158239
TX 2004 Brazoria 2	Bird 9473	2004	Brazoria Co., TX	Bluejay	V1	DQ158241
TX 2004 Swisher	Bird 9472	2004	Swisher Co., TX	House sparrow	V1	DQ158240
TX 2004 Randall	Bird 9477	2004	Randall Co., TX	House sparrow	V1	DQ158242
TX 2004 Galveston	Bird 9483	2004	Galveston Co., TX	Bluejay	V1	DQ158243

Isolates in bold have been completely sequenced.

P=RNA extracted from infected tissue

V1=RNA extracted from original isolation in Vero cells

V2=RNA extracted after single Vero cell passage

Table 2-3. Primers used for partial and complete genome sequencing.

PrM-E	
401+	Nonstructural protein genes
AAA AGA AAA GAG GAG GAA AG	
	4444+
1219-	GAT GAT GAT GGA AAT TTT C
GTT TGT CAT TGT GAG CTT CT	
	5417-
1101+	GGA GAC ATC AGC CTG
GAT GAA TAT GGA GGC GGT CA	
	5364+
1816-	TGA GAT CGT TGA TGT C
CCG ACG TCA ACT TGA CAG TG	
	6351-
1751+	CGT GAT GAC TTC AAC
TGC ATC AAG CTT TGG CTG GA	
	6269+
2504-	CAT ACC ATG ACC GGA AAT
TCT TGC CGG CTG ATG TCT AT	
	7282-
5' UTR	CCA TGT AAG CAT AGT GGC
1+	
AGT AGT TCG CCT GTG TGA	7087 +
	ACG TCA GAC TAC ATC AAC ACT T
533-	
CAG CAG CTG TTG GAA T	8060-
	ACT CCA CTC TTC ATG GTA A
Capsid	
WN 132 +	7999+
GAA AAC ATC AAG TAT GAG G	CAT GAA GAA CCA CAA CTG GT
WN 240 -	9043-
GAG GTT CTT CAA ACT CCA T	CCA TCA TGT TGT AGA TGC A
Nonstructural protein genes	8968+
2418+	TTT TGG GAG ATG GTG GAT GAG GAG
TGG AGG AGT TTT GCT CTT C	
	9804-
3238-	AAC CTG CTG CCA GTC ATA CCA CCC C
TGT ACC CTG GTC TCC TGT	
	9730+
3112+	AAT GCT ATG TCA AAG GTC C
GAA GTC AAA TCA TGC ACC	
	3' UTR
4037-	10660-
CTG TAC ACA TCA AGG TTT AAG	CCT GGG GCA CTA TCG
3849+	10460+
TTT CTT CCA AAT GGC TTA C	GCC ACC GGA AGT TGA GTA
4603-	10958-
CTC CTC TCT TTG TGT ACT GA	CCT GTG TTC TAG CAC CAC

Figure 2-1. Geographic location of isolates used in microevolution studies (Chapter 3).

Figure 2-2. Geographical location of Texas isolates used in microevolution studies (Chapter 3).

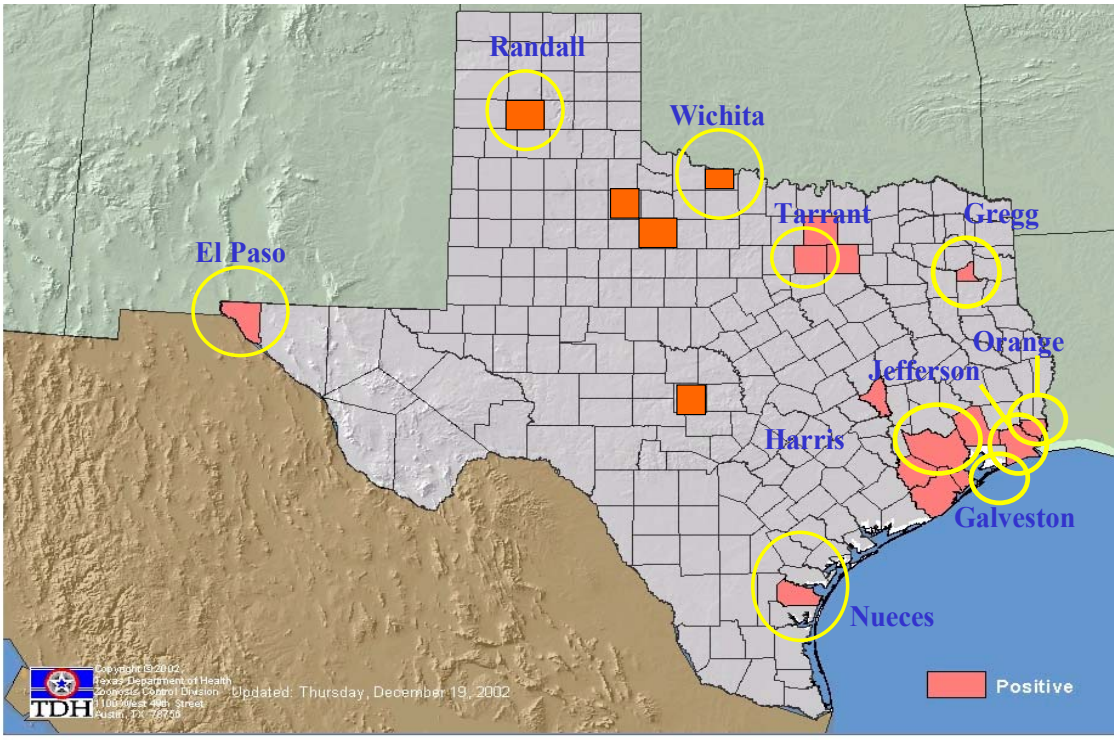


Figure 2-3. Geographic location of isolates used in North American phylogenetics analysis (Chapter 4).

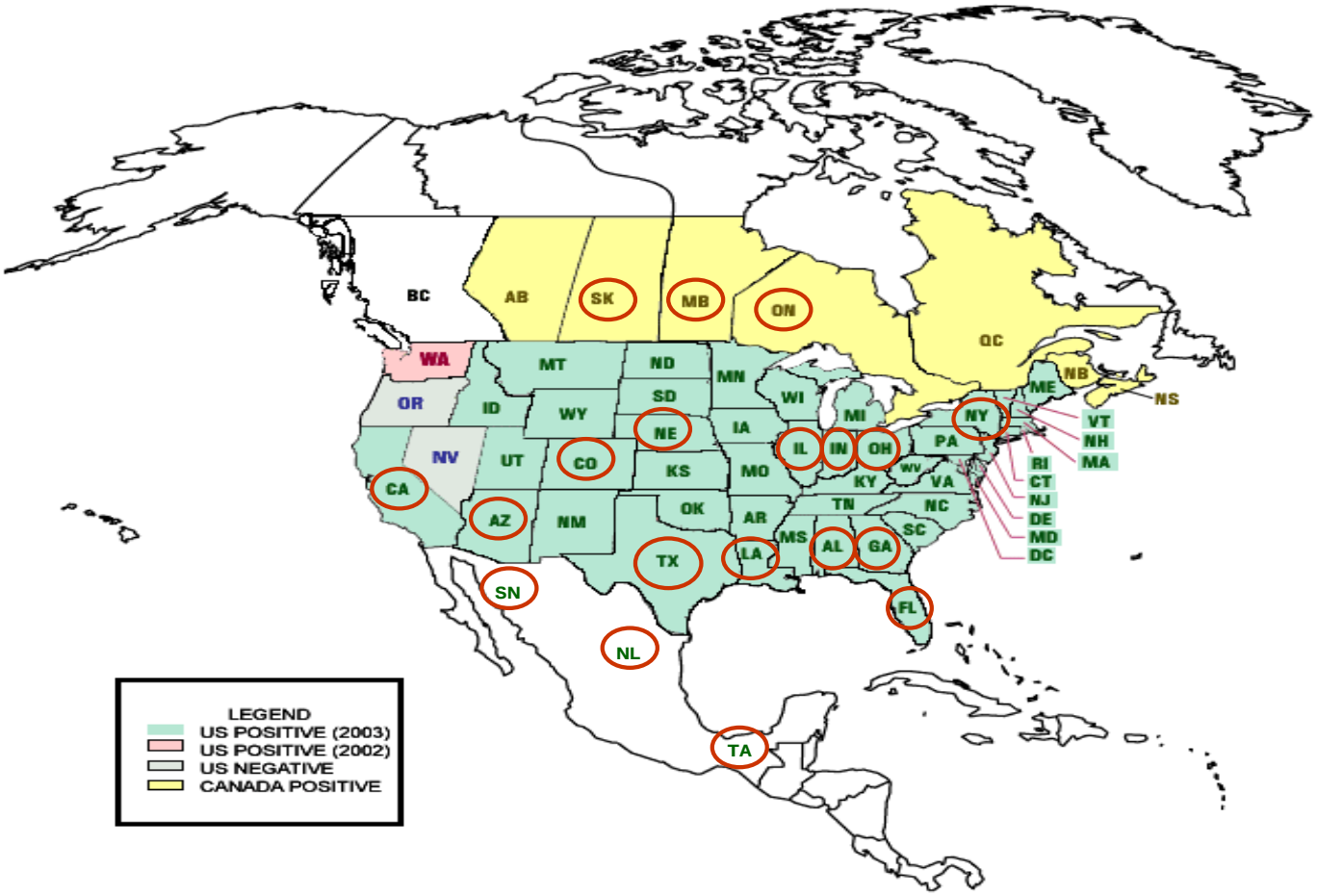
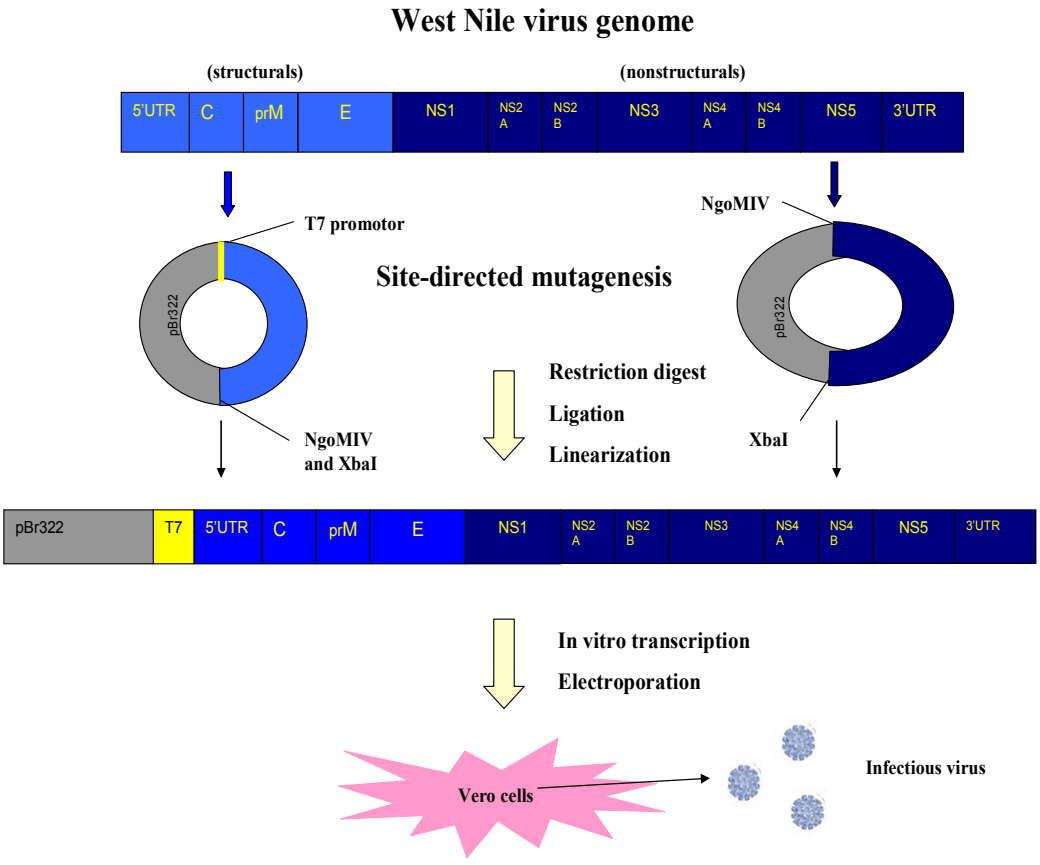


Figure 2-4. Cartoon of WN-NY99 infectious clone.



CHAPTER 3

MICROEVOLUTION OF WEST NILE VIRUS IN THE UNITED STATES, 2001-2002

Introduction

West Nile virus (WNV) is a member of the genus *Flavivirus*, family Flaviviridae and belongs to the Japanese encephalitis virus serocomplex. Until 1999, the geographical distribution of the virus was limited to Africa, the Middle East, India, and western and central Asia with occasional epidemics in Europe (Hayes, 1989; Murgue et al., 2002). As of December 2002, however, the distribution of the virus had expanded to include 44 states of the continental United States and southern regions of five Canadian provinces from Saskatchewan to Nova Scotia (CDC, 2002). Over the course of three years, the virus traversed North America presumably from New York City, where it was first isolated during the summer of 1999 (Anderson et al., 1999; Briese et al., 1999; Lanciotti et al., 1999; Steele et al., 2000). Partial nucleotide and complete genome sequence analysis of several WN strains isolated in the northeastern U.S. during 1999 and 2000 revealed that these isolates were most closely related to a WNV isolated from the brain of a dead goose in Israel in 1998 (Lanciotti et al., 1999; Jia et al., 1999; Lanciotti et al., 2002). Subsequent establishment of WNV across the eastern and mid-western regions of

North America from 1999 through 2001 set the stage for the rapid and widespread movement of the virus across the remainder of the continent during the summer of 2002, resulting in the highest number of annual case reports and deaths attributed to WNV in humans, equines, and birds documented since the discovery of the virus in North America. Surveillance programs initiated by public health agencies and research institutions, together with diagnostic laboratories, resulted in the collection of hundreds of WN isolates across the U.S. and Canada from various sources including mosquitoes, humans, equines, birds, as well as a number of other vertebrate species (CDC, 2002). Phylogenetic comparisons of partial and complete nucleotide sequences from isolates collected in the northeastern U.S. during 1999 and 2000 demonstrated a high degree of genetic similarity to the prototype New York strain, WN-NY99 (GenBank accession no. AF196835), with nucleotide identities of ≥99.8% and amino acid identities of ≥99.9% (Lanciotti et al., 2002; Anderson et al., 2001; Ebel et al., 2001; Huang et al., 2002). While these studies confirmed limited genetic divergence of northeastern isolates collected in 1999 and 2000 in comparison to WN-NY99, there was little published information describing the continuing divergence of WNV as its temporal and spatial distribution expanded (Beasley et al., 2003). In order to assess the extent to which WNV had evolved since its introduction in North America, the partial nucleotide and deduced amino acid sequences of WNV isolates collected during the summer and fall of 2001 and 2002 were analyzed and compared to a homologous sequence region of WN-NY99. Collaborations between the University of Texas Medical Branch (UTMB) and a number of public health

agencies from across the U.S. allowed for the collection of 22 isolates of WNV representing several geographically distinct regions within the U.S.

Results

Geographical location of WNV isolates and nucleotide sequences of the prM and E protein genes

Nucleotide sequences representing a 2004 nucleotide region of the complete prM-E protein genes of WNV (genomic nucleotides 466-2469) of the 18 isolates collected in 2001 and 2002, plus four of the previously published southeast Texas strains (Beasley et al., 2003), were compared to a homologous sequence region of the prototype WNV, WN-NY99 (See Table 3-1 for details of all isolates analyzed). Of the 22 isolates analyzed, 16 were collected from 10 different Texas counties, and two each from Illinois and Alabama, plus one each from Colorado and Louisiana. All isolates were from 2002 except two from Alabama in 2001 (Fig. 3-1). Sequence alignments comparing WN-NY99 to individual 2001 and 2002 isolates revealed up to 7 nucleotide mutations and 3 amino acid substitutions among the 22 isolates analyzed (Table 3-1 and Table 3-2). Nucleotide mutations occurred at 33 positions (9 in prM, 24 in E) with a total of 7 amino acid substitutions (2 in prM, 5 in E). The maximum nucleotide divergence of the 22 isolates from WN-NY99 was 0.35% with an average nucleotide divergence of 0.18%.

Phylogenetic relationships of WNV isolates

Several of the nucleotide mutations identified in this study were shared by many isolates (Table 3-1, Table 3-2 and Fig. 3-2). Two nucleotide mutations at residues 1442 (conservative amino acid substitution of Val to Ala at position E159) and 2466 were shared by 14 of the 22 isolates, with 10 of these 14 isolates sharing an additional non-coding nucleotide mutation at residue 660. Five different nucleotide mutations - at residues 969, 1192 (amino acid substitution of Thr to Ala at position E76), 1356, 2154, and 2400 - were shared by seven isolates, all of which were collected from coastal regions of southeast Texas. The isolate from Louisiana differed from WN-NY99 at only one nucleotide (residue 807) over the region studied and did not share any nucleotide mutations with other isolates from this study. In comparison, all the other nucleotide mutations identified in this study were not shared by nucleotide sequences reported previously from isolates collected in the northeastern U.S. during either 1999, or 2000, or 2001 (Lanciotti et al., 2002; Anderson et al., 2001; Ebel et al., 2001; Huang et al., 2002). Because these mutations were unique to isolates sequenced during this study, these results did not reveal a closer genetic relationship to isolates from either 2001, or 2000, or 1999. However, the two isolates in this study that were collected in 2001 (Alabama-1; Alabama-2) did share two nucleotide mutations (residues 1442 and 2466) with 12 of the other isolates collected in 2002. Construction of a phylogenetic tree by maximum parsimony analysis (Fig. 3-3) illustrates the genetic proximity of isolates from this study to those collected from the northeastern U.S. in 1999, 2000, and 2001. Branch groupings revealed both temporal and geographical separation of isolates, with those collected in

the northeastern U.S. in 1999, 2000, and 2001 representing a distinct clade relative to isolates collected in 2002. An exception to this grouping was an isolate from Louisiana collected in 2002, which grouped with northeastern U.S. isolates from 1999-2001. Significantly, WN viruses from the southeastern coast of Texas also comprise a clade of their own, separating these isolates from other 2001 and 2002 isolates collected from various regions within the U.S. It is significant to note that a WNV isolate collected from a Missouri dog in 2002 (GenBank accession no. AY160126) also shared a nucleotide mutation (residue 2466 C to U) with the 2002 isolates from this study. Although the entire prM-E protein genes of this isolate was not reported, it is likely that this isolate represents an additional member of the large 2002 clade.

Quasispecies

In a previous report by Beasley et al. (2003) concerning the genetic divergence of WNV since its introduction into the U.S., the authors described the existence of a quasispecies population within a single WNV isolate from Harris Co., TX. In order to determine if nucleotide mutations that define the southeast coastal Texas variant were uniform throughout the quasispecies population of a select isolate, the WN1751/WN2504A PCR product derived from WNV isolate Galveston Co., TX-3 was cloned into pGEM-T Easy. Ten clones were sequenced to obtain homologous regions of 700 nucleotides, which were then compared to the Galveston Co., TX-3 consensus sequence. This region contained the U to C mutation at nucleotide 2154 and the U to C

mutation at nucleotide 2400. Five of the ten clones were identical to the consensus sequence, while the other five clones each had one or two nucleotide changes from the consensus for a total of eight nucleotide changes (Table 3-3). None of the mutations identified represented amino acid substitutions, and unlike the 2001-2002 variant population (Beasley et al., 2003), none of the mutations encoded a stop codon. The maximum nucleotide divergence of individual clones was 0.28% (mean = 0.11%). Furthermore, none of the nucleotide changes identified in the five clones were shared with WNV strains representing the 2001-2002 variant, nor were there any nucleotide changes identified at two of the nucleotide positions that defined the southeastern coastal Texas variant. These results suggest that none of the virus genomes existing in a quasispecies population from WNV isolate Galveston Co., TX-3 contained nucleotide mutations characteristic of the 2001-2002 variant identified in this study.

Discussion

Phylogenetic comparisons of a 2004 nucleotide region encoding the entire pre-membrane and envelope protein genes (prM-E) of each isolate have revealed the most divergent variants of WNV in North America to date and provide evidence of the possible emergence of a dominant variant circulating in many regions of the U.S. Furthermore, these results indicate geographical clustering of distinct variants within and between states and reinforce previous evidence supporting the likelihood of multiple introductions of virus into the state of Texas (Beasley et al., 2003). Sequence

comparisons of a 2004 nucleotide region of 22 WNV isolates collected during the summer and fall of 2001 and 2002 revealed the highest degree of nucleotide divergence from WN-NY99 to date. Studies by Lanciotti et al. (2002) and Huang et al. (2002) have shown that the complete genomes of several WN isolates collected in 1999, 2000, and 2001 share ≥99.8% nucleotide identity to WN-NY99 with three or fewer amino acid substitutions in the entire polyprotein. Similar studies of partial nucleotide sequences conducted by Anderson et al. (2001) and Ebel et al. (2001) reported up to 3 nucleotide mutations over a region of 921 nucleotides and 1,503 nucleotides from isolates collected in Connecticut in 1999 and 2000 and New York in 2000, respectively. Although these studies have compared a larger portion of the genome than earlier studies of partial nucleotide sequences, we have identified individual isolates with as many as 7 nucleotide mutations and 3 amino acid substitutions with a maximum divergence of 0.35% from the homologous region of the prototype North American WNV, WN-NY99. The nucleotide mutations identified in this study were not shared by previously sequenced isolates from 1999, 2000, or 2001 (Lanciotti et al., 2002; Anderson et al., 2001; Ebel et al., 2001; Huang et al., 2002), and represent new nucleotide changes in the North American WNV population. Since these changes were not shared with other previously reported WNV sequences, the isolates analyzed in this study do not show a greater genetic similarity to northeastern isolates from either 1999, 2000, or 2001. It is important to note, however, that several of these nucleotide changes (660, 969, 1356, 2154, 2400, and 2466) are observed in other Old World WNV strains from both lineage 1and 2 (Table 3-4). Each of these changes represents a non-coding mutation from either a C to U or U to C in the

third codon of the open reading frame; nucleotides at these positions may revert back to nucleotides observed in the more ancestral Old World strains.

These results also suggest the geographical clustering of genetically distinct variants. Seven of the 22 isolates, all of which were collected from coastal regions of southeast Texas, share 5 nucleotide mutations unique to only these isolates. Fourteen of the other isolates, which represent the CDC defined East South Central (AL), West South Central (LA and TX), East North Central (IL), and Mountain (CO) regions (CDC, 2002), all share 2 unique nucleotide mutations not identified in other isolates (Fig. 3-2). The results of this study support those of Beasley et al. (2003), which suggest that during the summer of 2002 there were at least two separate introductions of WNV into the state of Texas. It is possible that these results reflect the unique migratory patterns of North American birds acting as reservoir hosts for WNV. As has been illustrated by Rappole et al. (2000), many North American birds follow well-documented migration routes from summer grounds in the northeastern U.S. to southern areas that are classified as the southeastern U.S., circum-Gulf, trans-Gulf, and Caribbean/western North Atlantic routes. For example, the Laughing Gull (*Larus atricilla*) has been known to follow a circum-Gulf route as it travels from the northeastern U.S. to stop-over sites along the northern and western Gulf coast on its way to Mexico or Central America. Because certain species of birds show a more limited geographical range than others, it is possible that geographically clustered populations of distinct genetic variants, as is observed with isolates collected from coastal regions of southeast Texas, arise as a result of restricted migratory routes of a given species of bird. This hypothesis is supported by a number of

studies. Peiris and Amerasinghe (1994) have identified a group of geographically restricted antigenic variants of WNV confined to southern India. Because of the lack of bird migratory routes linking southern India with the Middle East and Africa, a distinct antigenic group exists exclusively in southern India. Furthermore, numerous studies have shown antigenic variation among WNV strains that correlate to geographically distinct regions and restricted migratory patterns of birds (Hammam et al., 1965; Price and O'Leary, 1967). Phylogenetic comparisons of Indian viruses to other WN viruses reveal similar findings, which place Indian WN strains in a unique clade of the lineage 1 WN viruses (Lanciotti et al., 2002; Burt et al., 2002). Recent studies from Israel by Malkinson et al. (2002) also support the role of migratory birds in the dispersion of unique WNV variants in geographically distinct regions. The results of this study also support an alternative hypothesis that explains the continental spread of WNV as a consequence of transmission between local bird and mosquito populations in a given region. This mechanism allows for spread of the virus from region to region over shorter distances, in contrast to the long distances traveled by migratory birds (Rappole and Hubalek, 2003). The findings of a dominant variant that exists over a large part of the U.S., together with evidence of a geographically distinct southeast coastal Texas variant, suggest that both mechanisms of spread have influenced the genetic distribution and spread of WNV in the U.S.

There is little genetic evidence to support or refute the hypothesis of WNV becoming established in an enzootic transmission cycle in a particular geographic area compared to reintroduction of WNV into a particular area each year when the transmission season

begins. Similarly, due to the limited published data detailing the year to year genetic changes observed in WNV, it remains to be established whether or not the virus is becoming endemic in particular regions of the U.S. This question will be answered in part by determining baseline phylogenetic results of specific variants in a geographic area and by analyzing isolates collected in sequential transmission seasons.

Although the isolates analyzed in this study do not represent the entire temporal and geographic distribution of WNV in North America, it appears that at least some nucleotide mutations have been conserved among West Nile viruses now circulating across the continent. If indeed the conservation of these mutations is the result of selective pressure, such as the continued capacity to replicate in both arthropod and vertebrate hosts, rather than random mutations occurring as a consequence of genetic drift, one would expect these mutations to be conserved in virus isolates collected in other regions of North America. Further investigation concerning the genetic composition of viruses from additional regions of North America will define the extent to which dominant variants have emerged. If dominant variants do continue to emerge across the U.S., phylogenetic analyses will help researchers to monitor the spread of WNV in North America and may provide explanations for the rapid and widespread movement of this newly emerging virus in North America. Similarly, identification of the genetic composition of WNV isolates from other regions of the U.S. and Canada, as well as comparisons of these isolates to isolates collected in 2003, will continue to define evolutionary relationships of WN viruses circulating in North America and facilitate predictions concerning the primary mechanisms of transmission and spread of the virus.

Table 3-1. Nucleotide mutations in sequences of the prM protein gene of 22 West Nile virus isolates obtained during 2001 and 2002 compared to WN-NY99.

				prM (501 ntds)								
Strain	Source	Collected	RNA origin	491 (prM9)	507	549	621	660	679 (prM72)	690	807	903
WN-NY99 (AF196835)	Flamingo	6/1/1999	Brain	A (Lys)	A	U	A	C	U (Ser)	C	C	G
Harris Co., TX (AY185906)	Bluejay	6/11/2002	Brain/Vero					U				
Harris Co., TX (AY185907)	Bluejay	6/10/2002	Brain/Vero					U				
Nueces Co, TX - 1	C. quinquefasciatus	8/6/2002	BHK	G (Arg)				U				
Nueces Co., TX - 2	C. quinquefasciatus	9/17/2002	BHK					U				
Gregg Co., TX	C. quinquefasciatus	9/25/2002	BHK					U				
Tarrant Co., TX	C. restuans	9/30/2002	Vero									
Wichita Co., TX	C. quinquefasciatus	10/23/2002	BHK					U				
Randall Co., TX	C. tarsalis	9/18/2002	BHK					U				
El Paso Co., TX	C. tarsalis	8/26/2002	Vero					U				
Illinois - 1	C. pipiens	8/2/2002	Vero			A				U		
Illinois - 2	Dog	8/1/2002	Kidney/Vero		G			U				
Alabama - 1	Crow	10/5/2001	Brain/Vero									
Alabama -2	Crow	9/10/2001	Brain/Vero									
Colorado	Red-tailed Hawk	8/1/2002	Brain/Vero					U				
Louisiana	C. salinarius	8/6/2002	Vero								U	
Galveston Co., TX-1 (AY185914)	Bluejay	8/2/2002	Vero				G		A (Thr)			
Galveston Co., TX-2 (AY185913)	Bluejay	7/19/2002	Vero						A			
Galveston Co., TX-3	C. quinquefasciatus	8/21/2002	Vero									
Jefferson Co., TX - 1	C. quinquefasciatus	8/6/2002	BHK									
Jefferson Co., TX - 2	C. quinquefasciatus	7/2/2002	BHK									A
Jefferson Co., TX - 3	Human	8/24/2002	CSF/Vero									
Orange Co., TX	C. quinquefasciatus	7/3/2002	Vero									A

Amino acids in brackets refer to amino acid substitutions encoded by nucleotide change.

Table 3-2. Nucleotide mutations in sequences of the E genes of 22 West Nile virus isolates obtained during 2001 and 2002 compared to WN-NY99.

	Envelope (1503 ntds)																							
Strain	969	1038	1065	1071	1118 (E51)	1137	1179 (E71)	1192 (E76)	1293	1356	1377	1442 (E159)	1443	1554	1557	1581	1728	1830	2094	2154	2190	2392 (E476)	2400	2466
WN-NY99	C	U	C	U	C (Ala)	C	A (Lys)	A (Thr)	C	C	C	U (Val)	U	U	C	U	A	U	A	U	U	G (Ala)	U	C
Harris Co., TX												C (Ala)										A (Thr)		U
Harris Co., TX			U			U						C												U
Nueces Co, TX - 1											U	C									C			U
Nueces Co., TX - 2					U (Val)				U		U	C												U
Gregg Co., TX												C	C											U
Tarrant Co., TX		C										C												U
Wichita Co., TX												C												U
Randall Co., TX					U		C (Asn)					C												U
El Paso Co., TX												C												U
Illinois - 1												C												U
Illinois - 2												C			U	C		C						U
Alabama - 1												C												U
Alabama -2												C					U							U
Colorado												C							G					U
Louisiana																								
Galveston Co., TX-1	U							G (Ala)		U										C			C	
Galveston Co., TX-2	U							G		U										C			C	
Galveston Co., TX-3	U							G		U				C						C			C	
Jefferson Co., TX - 1	U							G		U										C			C	
Jefferson Co., TX - 2	U			C				G		U										C			C	
Jefferson Co., TX - 3	U							G		U										C			C	
Orange Co., TX	U			C				G		U										C			C	

Amino acids in brackets refer to amino acid substitutions encoded by nucleotide change.

Table 3-3. Nucleotides and their deduced amino acids that varied among individual clone sequences of a fragment of the E protein gene (residues 1769-2496) of the WNV Galveston Co., TX-3

	Nucleotide							
Clone	1779	1787	1798	1871	2162	2168	2232	2469
Consensus	U	U	A	A	A	G	A	U
1			G				G	
2	C	C						
4				G				C
6					G			
7						A		
Clones 3, 5, 8, 9, 10 were identical to consensus sequence.								

Table 3-4. Nucleotide changes from WN-NY99 observed in 2001 and 2002 WNV isolates that are conserved in Old World WNV isolates with complete genomes available from GenBank.

660 (C to U)	969 (C to U)	1356 (C to U)	2154 (U to C)	2400 (U to C)	2466 (C to U)
WN Uganda 1937 [a](M12294) WN LEIV-Krnd88-190 (AY277251)	WN IS-98 STD (AF481864) WN Eg101 (AF260968) WN Ast99-901 (AY278441) WN LEIV-Krnd88-	WN Eg101	WN Uganda 1937	WN Uganda 1937 WN LEIV-Krnd88-190 WN Eg101 WN Ast99-901 WN RO97-50 (AF260969) WN VLG-4 (AF317203) WN KN3829 (AY262283) WN Italy 1998-equine (AF404757) WN LEIV-Vlg00-27924 (AY278442) WN VLG-4	WN Uganda 1937

[a]Number in parenthesis represents GenBank accession no.

Figure 3-1. Locations of collected isolates, 2001-2002.

Fig. 3-2. Phylogram based on maximum parsimony analysis comparing a 2,004-nucleotide sequence of WN-NY99 (GenBank accession no. AF196835) with 22 WNV isolates collected during 2001 and 2002.

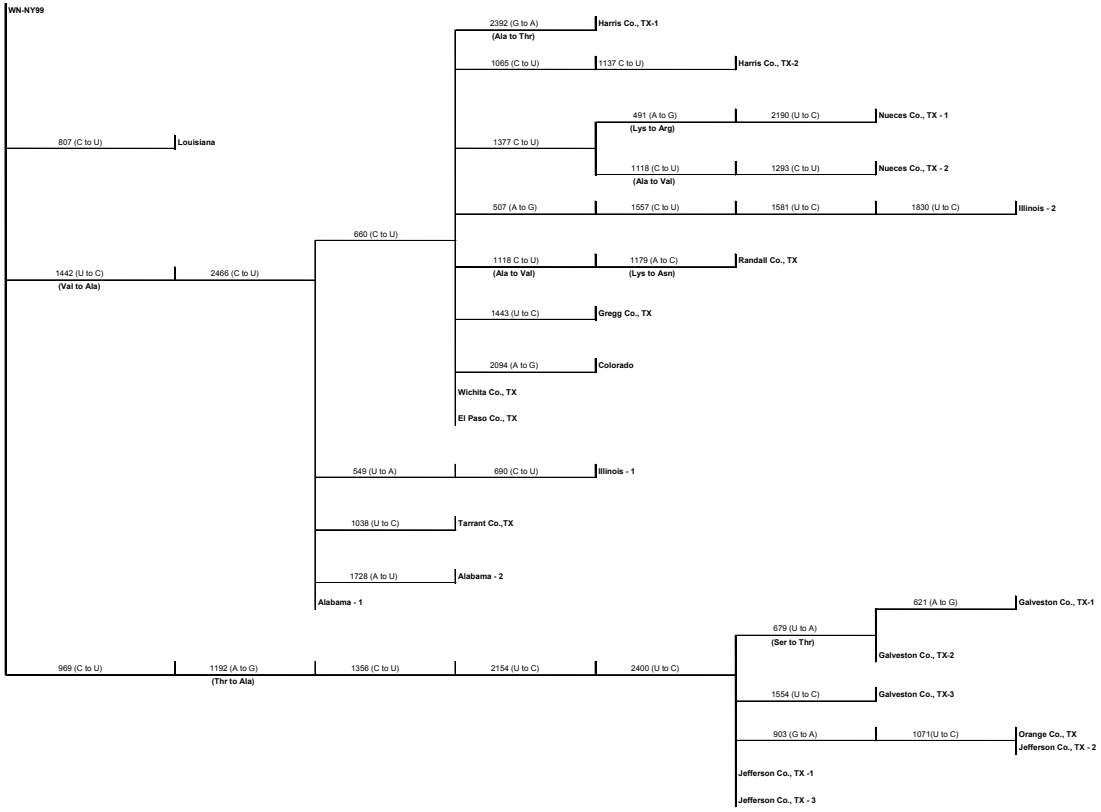
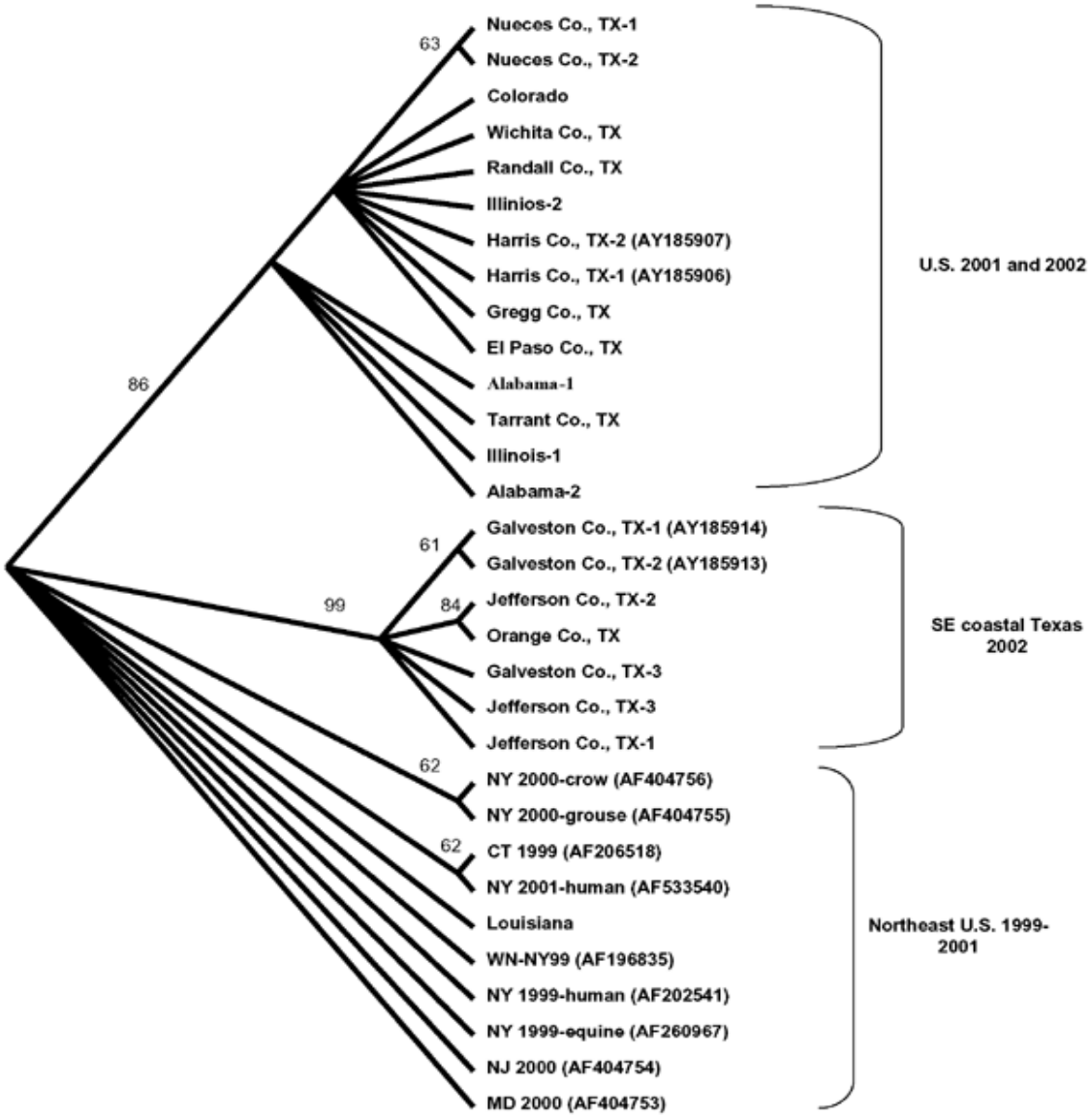


Fig. 3-3. Cladogram based on maximum parsimony analysis comparing a 2,004-nucleotide sequence of 22 West Nile virus isolates collected during 2001 and 2002 with a homologous region of WN virus isolates collected in 1999, 2000, and 2001 from the northeastern United States. Numbers indicate bootstrap confidence estimates based on 500 replicates for clades supported to the right. Numbers in parenthesis represent GenBank accession numbers.



CHAPTER 4

PHYLOGENETIC ANALYSIS OF NORTH AMERICAN WEST NILE VIRUS ISOLATES, 2001-2004: EVIDENCE FOR THE EMERGENCE OF A DOMINANT GENOTYPE

Introduction

West Nile virus (WNV) (Family Flaviviridae: Genus *Flavivirus*) has recently undergone a dramatic increase in range to include much of North America. Since the summer of 1999, the distribution of WNV has expanded to include the contiguous U.S. states and seven Canadian provinces, as well as Mexico and the Caribbean Islands (Estrada-Franco et al., 2003; Komar et al., 2003; Blitvich et al., 2003; Dupuis et al., 2003; Quirin et al., 2004) and Colombia (Kramer pers. comm.). Because of its relatively recent introduction, studies concerning the evolution of WNV are important for understanding the extent to which the virus has mutated as its temporal and geographic distribution have expanded and to recognize the accumulation of mutations in the genome that may be important to transmission and perpetuation of this virus in nature. Nucleic acid sequencing of WNV isolates collected across the U.S. since 1999 have identified

mutations to the genome when compared to the prototype New York strain, WN-NY99 (GenBank accession no. AF196835) (See Chapter 3). These mutations reveal the presence of distinct genetic variants that group in a temporally and geographically dependent manner (Lanciotti et al. 2002; Beasley et al., 2003; Davis et al., 2003; Ebel et al., 2004) and suggest the emergence of a dominant genetic variant that is now distributed across much of North America. Additional genetic and phenotypic studies suggest that the increased isolation frequency of the dominant genetic variant during 2002 and 2003 may be due to the possibility of enhanced mosquito transmission efficiency of this variant in *Culex pipiens*, one of the main vectors in the northeastern United States (Ebel et al., 2004). In order to assess the genetic divergence of WNV as its spatial and temporal distribution have expanded and to identify mutations in the WNV genome that may be conserved in the North American WNV population, both partial and complete genome sequences of WNV isolates collected from across North America were characterized. The prM and E protein genes of 43 WNV isolates collected during 2001-2004 from the U.S., Canada, and Mexico were sequenced and compared to other North American WNV sequences available in GenBank and from collaborators. In addition, the complete genomes of 5 isolates were sequenced in order to attain a more robust comparison of the genetic relationships of WNV isolates because of the relatively few phylogenetically informative sites in the prM and E protein genes of isolates studied. Sequence alignments were constructed in order to identify nucleotide and amino acid mutations in the WNV genome, which may support evidence that a dominant genetic variant has emerged throughout North America. The individual genes of the WNV genome were compared to

test whether or not certain genes may be more prone to the accumulation of mutations, and thus may act as surrogates to complete genome sequencing when performing phylogenetic analyses. Phylogenetic trees were generated using both individual viral gene and complete genomic sequences in order to compare the relationships among WNV isolates generated using each method, and to identify those genes of the WNV genome that are the most parsimony informative in the phylogenetic analysis of North American isolates. Collectively, the data suggest that there has been continued divergence of WNV as the temporal and spatial distribution of the virus has expanded.

Results

Nucleotide sequencing of North American WNV isolates

The prM and E protein genes of 43 isolates were sequenced and deposited in GenBank with information pertaining to source, collection date, and location of the isolate (Accession No. DQ158207-DQ158250) and are shown in italics in Table 4-1. The prM and E genes of 31 additional isolates were sequenced by other laboratories and provided for this study (Brault, Ebel, Lanciotti personal communication). Complete genomes of 25 additional isolates (5 from the author shown in italics in Table 4-1 with additional genomes courtesy of Ebel and Lanciotti) were sequenced and also submitted to GenBank (Accession No. DQ164186-DQ164206). Table 4-2 shows the percentage nucleotide and deduced amino acid sequence divergence among the complete genomes and ORFs of all newly sequenced WNV isolates compared to prototype NY99. All North

American West Nile viruses maintain a high degree of both nucleotide and amino acid sequence conservation over the complete genome with an average degree of nucleotide and amino acid divergence from NY99 of 0.24% and 0.09%, respectively. In comparison to NY99, both the nucleotide and amino acid divergence tended to increase as the year of isolation increased with the exception of a few isolates (e.g., TX 2004 Harris 4). The highest degree of nucleotide divergence from NY99 was 0.39% for AZ 2004 while the degree of amino acid sequence divergence was highest for NY 2003 Rockland, which had seven amino acid substitutions from NY99. Overall, the highest degree of both nucleotide and amino acid divergence was between AZ 2004 and NY 2003 Rockland at 0.58% and 0.38%, respectively. Interestingly, the degree of nucleotide and amino acid divergence did not always correlate for a given isolate because the majority of the nucleotide mutations were synonymous. Each of the 25 isolates that were completely sequenced had at least one amino acid substitution relative to the consensus sequence derived from an alignment of all completely sequenced isolates (Table 4-3). In general, isolates made during 2003 and 2004 had accumulated a larger number of amino acid substitutions than isolates made in 2001 and 2002. Table 4-3 shows that the majority of the isolates share a conserved amino acid substitution at position E159 from a Val to Ala with the exception of some isolates from New York, 2001-2003, and an isolate from Ohio, 2002. Several amino acid substitutions were also shared by isolates made in the same state, though not necessarily the same year. Deduced amino acid substitutions were identified in each of the 25 isolates with the largest number of amino acid substitutions occurring in NS5 (4.6%).

Phylogenetic analysis of WNV isolates based on the prM and E protein genes

Several phylogenetic trees were generated from prM and E protein genes and 5' and 3' untranslated regions, as well as complete genomes in order to ascertain the phylogenetic relationships of the isolates studied. RT-PCR was used to amplify and subsequently sequence a 2004 nucleotide region of the prM and E protein genes of 74 WNV isolates to broaden the scope of the phylogenetic analysis because there are a larger number of partial sequences (primarily the prM and E protein genes) of WNV isolates available in GenBank. A Bayesian analysis was used to generate a consensus tree based on the prM and E protein gene alignment of these 74 North American WNV isolates in comparison to all other North American WNV isolates for which the prM and E sequence was available in GenBank (a total of 108 isolates were compared) (Figure 4-1). Bayesian analysis was used to analyze the prM and E protein genes because it was found to be a more efficient method for the generation of equally or more robust phylogenetic trees in comparison to more commonly used phylogenetic methods (i.e., parsimony, likelihood analysis). This method proved especially useful in generating consensus trees rapidly when comparing isolates that differed by only a few informative sites. The tree was rooted with the most closely related Old World WNV, Israel-1998 (Accession no. AY033389), in order to generate a tree illustrating more parsimony informative sites. This analysis revealed three distinct clades comprised of North American isolates. The apical clade was comprised of isolates collected from all over North America that were collected during 2002-2004 and was termed the "North America 2002-2004" clade. Within this clade were many subclades, which consisted of isolates that were highly

conserved at the nucleotide sequence level and some that were identical to one another. Most of the isolates within these subclades were made in the same state and during the same year, although there was evidence that some isolates made in close geographic proximity tended to cluster into subclades. Basal to the North American 2002-2004 clade was a sister clade that represents an intermediate grouping of 13 isolates that shared some degree of sequence identity with the North American 2002-2004 clade. This sister clade was made up primarily of isolates from eastern regions of the U.S., including New York, Alabama, Indiana, Illinois, Louisiana, and eastern Texas. An isolate from Tabasco State, Mexico made up a second sister clade because of a conserved nucleotide substitution at position 2466 in the envelope protein gene that was also found in all apical isolates. Two additional Mexican isolates from Sonora State and Nuevo Leon State did not group with the isolate from Tabasco State, which again suggested the occurrence of multiple introductions of WNV into Mexico as has been suggested by others (Estrada-Franco et al., 2003; Blitvich et al., 2004). Interestingly, both of the Mexican isolates fall into the most apical clade, which also contains isolates collected from Arizona in 2003 and 2004, several isolates collected in Texas in 2004, and an isolate from Colorado in 2003. A second clade, termed the "Eastern U.S." clade, was comprised of isolates collected from the eastern U.S. states. The majority of these isolates were made prior to 2003, with the exception of two isolates from New York during 2003. This clade also contained isolates from the northeastern U.S. that were made during the early stages of the WNV epidemic, including isolates from New York in 1999. Although genetically more distantly related to Israel-1998 than the Eastern U.S. isolates, the most basally positioned clade was

comprised of "Southeast coastal Texas" isolates from 2002. The basal position of this clade occurred as a result of several non-parsimony informative nucleotide mutations shared by isolates in the Eastern U.S. clade and the North American 2002-2004 clade, none of which were found in the southeast coastal Texas 2002 clade. Interestingly, isolates made from the southeast coast of Texas in 2003 and 2004 belonged to the North American 2002-2004 clade suggesting that the southeast coastal Texas 2002 genotype has become extinct or displaced in this particular region.

Phylogenetic analysis of WNV isolates based on the complete genomes

To further define the phylogenetic relationships among North American WNV isolates and to confirm the relationships as delineated by the prM and E protein gene sequences, the complete genomic sequencing of 5 isolates was undertaken together with genomic sequences of 20 isolates from collaborators to provide a more robust comparison. A phylogenetic tree was generated by maximum likelihood analysis to show the relationship between the newly sequenced isolates and those North American isolates available in GenBank. A total of 36 North American WNV isolates from this and previous studies group in a monophyletic clade of the Lineage 1 WNV strains, which sits just apical to the closest related Old World strain, Israel-1998 (Fig. 4-2). The complete genome analysis generated a tree that also illustrates the presence of the North American 2002-2004 clade, relative to the Eastern U.S. clade, and illustrates the presence of more highly resolved subclades within each of the larger clades. As was noted in the prM and

E protein gene tree, the majority of fully resolved subclades were comprised of isolates that were made in the same geographic region and during the same year. It is interesting to note that the complete genome analysis placed an isolate from the southeast coast of Texas in the same clade as other isolates from the eastern U.S. Because of the lack of nucleotide mutations in isolates belonging to the Eastern U.S. clade relative to the North American 2002-2004 clade, sub-groupings were less resolved and the clade consisted of a large polytomy. Interestingly, the resolved sub-groupings were comprised of northeastern isolates that were made primarily in 2002 and 2003. Also of interest was an intermediate clade found between the two larger clades that is comprised of an isolate from Tabasco State, Mexico, an isolate from Ohio, 2002, and an isolate from New York, 2000, indicating the likelihood that additional clades within the North American monophyletic clade will continue to be identified as additional genome sequences become available.

Figure 4-3 shows two additional phylogenetic trees that represent the overall tree topology generated from analyses of each of the individual genes and untranslated regions of the WNV genome. Bootstrap values are shown only at clades and subclades that were resolved in each of the trees generated by analysis of individual genes or regions. The “group B” topologies are similar to those seen in the complete genome analysis (Fig. 4-2) in that the two large clades can be readily distinguished from one another with average bootstrap values > 85. The group B topology was generated from analyses of the prM, E, NS2A, NS3, NS4B, NS5, and 3’UTR. In contrast, “group A” trees (generated from analyses of the 5’UTR, capsid, NS1, NS2B, and NS4A) do not delineate between the two clades, and illustrate only the monophyletic North American

clade relative to Old World isolates. Also, group A trees did not place Israel-1998 as a distinct outlier to the North American clade. A lack of parsimony informative mutations within certain genes or untranslated regions is likely to relate to the lack of topology seen in the group A trees. Accordingly, the complete genome alignment of the completely sequenced WNV isolates was used to identify the nucleotide and deduced amino acid mutations that are conserved in isolates belonging to the North American 2002-2004 clade. Table 4-4 shows that no conserved mutations have accumulated in the genes (capsid, NS1, NS2B, and NS4A) or untranslated region (5'UTR) that produce trees with no resolution of clade topology. In contrast, at least one nucleotide mutation was conserved in individual genes or the 3' untranslated region of group "B" isolates resulting in trees that show the presence of multiple clades (Fig. 4-3). The largest number of conserved mutations was found in the NS3 gene. Despite the occurrence of 13 nucleotide mutations [nucleotides prM 660(C to U), E 1442(U to C), E 2466(C to U), NS2A 3774(U to C), NS2A 4146(A to G), NS3 4803(C to U), NS3 6138(C to U), NS3 6238(C to U), NS3 6426(C to U), NS4B 6996(C to U), NS5 7938(U to C), NS5 9352(C to U), 3'UTR 10851(A to G)]in each of the isolates belonging to the North American 2002-2004 clade, there has been only a single amino acid substitution conserved in these isolates (E V159A).

Discussion

Since the detection of WNV in the U.S. in 1999, studies concerning the molecular epidemiology of the virus have proven useful for tracing the geographic and temporal

spread, and interhost transmission of the virus (Davis et al., 2003; Ebel et al., 2004; Solomon et al., 2003; Estrada et al., 2003; Blitvich et al., 2004) and in detecting phenotypic variants (Beasley et al., 2004; Davis et al., 2004; Ebel et al., 2004). Because the location and year of the first virus isolations have been well documented (Lanciotti et al., 1999; Lanciotti et al., 2002), it has been possible to make genetic and phenotypic comparisons between isolates made at the onset of the North American WNV epizootic and isolates made from all subsequent years and over a broad geographic distribution (Anderson et al., 2001; Beasley et al., 2003; Davis et al., 2003, 2004; Ebel et al., 2000, 2004). Consequently, genome sequences of isolates made in 1999 and 2000 provide a genetic baseline allowing for the identification of novel mutations in the genomes of more recently isolated strains, which can then be used to infer phylogenetic relationships among isolates. Previous phylogenetic studies have focused on sequencing complete genomes of virus isolates from the northeastern U.S. and only partial sequences (primarily the prM and/or E protein genes) of isolates from other regions of North America. Accordingly, this study has utilized both partial and complete genome sequences of isolates from 2001-2004 in order to further characterize the phylogenetic relationships of North American WNV isolates and to better understand how this virus has evolved since its introduction six years ago.

The prM and E protein genes for 43 North American isolates were sequenced by the author and sequences for an additional 31 isolates were provided by collaborators. The prM and E protein genes were selected because previous studies of WNV molecular epidemiology have relied on sequencing of these regions. Also, because of the relatively

few complete nucleotide sequences of North American WNV isolates available for analysis prior to this study, the complete genomes of isolates representing a wide range of host species, collection years, and locations were either sequenced by the author or by collaborators. Although the prM and E protein gene sequences demonstrate an accurate and reliable phylogenetic representation, analysis of complete genomes of North American WNV isolates provided a more robust comparison of the evolutionary relationships between isolates and revealed additional mutations to the WNV genome that may have implications for the phenotypic properties that these isolates display (Davis et al., 2004; Beasley et al., 2004; Ebel et al., 2004). Moreover, complete genome analysis allows one to compare isolates at the individual gene level in order to identify those genes that may be more reliable for the comparison of parsimony informative sites among isolates. Therefore, this study has used a variety of methods to infer phylogenetic relationships of North American WNV isolates and to further define the evolution of WNV in North America. Data from complete nucleotide and amino acid sequence analyses supports previous findings (Davis et al., 2003; Beasley et al., 2003; Ebel et al., 2004; Lanciotti et al., 2002) suggesting that WNV has continued to diverge from progenitor isolates (those made in the northeastern U.S. during the early stages of the WNV epizootic) as the temporal and geographic distribution of the virus has expanded. It was also evident from this analysis that isolates displaying the highest degree of nucleotide and amino acid sequence divergence were collected, in general, after 2002. This may be significant as 2002 was the year of the largest recorded epidemic of arboviral encephalitis in North America and saw a vast geographic expansion by the virus

(CDC, 2002). One would expect that the nucleotide divergence would reflect the nature of the error-prone RNA dependent RNA polymerase of WNV in that as the virus continues to replicate over time and geographic location, additional mutations will continue to arise in the genomes of viruses. However, the results in Figure 4-3 show that nucleotide mutations throughout the genome are not random and that some genes have not accumulated any mutations. Although the majority of these mutations are silent, it is apparent from this and other studies that amino substitutions are not uncommon and have resulted in phenotypic variation within the North American WNV population (Beasley et al., 2004; Davis et al., 2004; Ebel et al., 2004). Despite the continuing divergence of the virus from progenitors, many of the accumulated nucleotide mutations in the genome have become fixed in currently circulating viruses. Many of these fixed mutations were found in isolates from Georgia, Texas, and Indiana in 2002 indicating their presence since at least 2002 and were most likely found in those viruses that were circulating at the forefront of the western expansion of the virus. Consequently, all isolates sequenced from the western U.S. contain the majority of those fixed mutations shown in Table 4-4, a result that may reflect the rapid westward progression of the virus from 2002 to 2004. In contrast, there has been only a single, conserved amino acid substitution (E V159A) shared in the majority of isolates sequenced in this study, perhaps reflecting the inability of the virus population to sustain substitutions in viral proteins that may be disadvantageous to viral fitness. Incidentally, this particular substitution is found in many Old World WNV strains and may represent a non-critical residue that has reverted back to a residue common in ancestral strains (Ebel et al., 2004). Undoubtedly, as the virus

continues to evolve additional amino acid substitutions will become fixed in the population at large, the consequences of which remain speculative at this point.

Phylogenetic analysis of prM and E protein gene sequences revealed three distinct clades within the North American WNV monophyletic group. The majority of isolates collected since 2002 belong to a single clade that has been termed the North American 2002-2004 clade. The emergence of this clade appears to correspond with the increased intensity of transmission and western progression of WNV across the U.S. during 2002 and indicates that as the distribution of the virus expanded, mutations in the genomes of circulating viruses began to accumulate and became fixed in the overall virus population. The finding that the majority of isolates made from 2002-2004 belong to this clade suggests the emergence of a dominant genetic variant whose distribution is now uniform across all of North America. It is also apparent that this variant has begun to replace other genetic variants in certain regions (i.e., Eastern U.S. and Southeast coastal Texas variants). A study by Ebel et al. (2004) suggests that this displacement and rapid spread may be correlated with increased transmission efficiency of the dominant genetic variant in *Culex spp.* mosquitoes when compared to other variants. Additional phenotypic studies will be required to characterize any fitness advantages that this variant may have over others.

Analysis of the prM and E genes of North American WNV isolates also separated many isolates of the North American 2002-2004 clade into subclades that represent isolates from certain regions of North America. Of note was a subclade consisting of isolates from Mexico, Arizona, and Texas suggesting that transmission of viruses in these

regions may be the result of localized virus spread by resident birds rather than by migratory birds that may introduce more distantly related viruses into a particular area. In many cases, several subclades were made up of isolates from the same state during the same year, illustrating that genetic distance among WNV isolates often reflects geographic and temporal distances. In contrast, some intermediate clades, as illustrated in the complete genome tree (Fig. 4-2), may also represent the long-distance spread of WNV by migratory birds, in contrast to the more localized spread by non-migratory species. Evidence that migratory birds are capable of the dispersal of discrete genetic variants over long distances is supported by the close genetic relatedness of the Tabasco 2003 isolate to a 2002 Ohio isolate.

As noted by Blitvich et al. (2004), Bayesian analysis of WNV isolates that differ by only a few informative sites can be used as a more efficient method for the generation of equally or more robust phylogenetic trees in comparison to more commonly used phylogenetic methods (i.e., parsimony, likelihood analysis). Similar topologies were created by the author using maximum parsimony, neighbor-joining, and Bayesian methods, but the Bayesian analysis generated equally or more significant confidence values at internal nodes. Bayesian analysis was preferred to other methods because the analysis was also less computationally exhaustive due to the large number of isolates compared in the prM and E protein gene sequence alignment. While analysis using complete genome sequences is clearly the most informative method of describing evolutionary relationships among viral isolates, a lack of complete sequences in public databases makes this technique of limited use. Thus, at the present time, Bayesian

analysis appears to be the most efficient method for describing the relationships between WNV isolates that share high nucleotide identities with few informative sites.

Complete genomic sequence analysis revealed similar tree topology as that found in the analysis of prM and E protein genes (Fig. 4-2 and Fig. 4-1, respectively). In comparison to Old World isolates, the North American monophyletic group remained divided into two separate clades, the North American 2002-2004 and Eastern U.S. clades. It is likely that the Southeast coastal Texas clade was not resolved by the complete genome analysis because only a single isolate from this clade has been completely sequenced (Granwehr et al., 2004). The results demonstrate that all isolates collected after 2002 with the exception of a few isolates from New York in 2003 are contained in the North American 2002-2004 clade. Thus, it is evident from the complete genome analysis that a dominant WNV variant has emerged in all of the regions sampled in this study. This analysis also reveals that isolates collected in close spatial and temporal proximity cluster in discrete subclades. Despite the existence of a large polytomy in the Eastern U.S. clade, the more highly evolved North American clade contains more fully resolved subclades, again suggesting that isolates collected outside of the northeastern U.S. and after 2002 depict the continuing divergence of isolates from progenitors and that their phylogenetic relationships can be better resolved because of a larger number of conserved mutations in their genomes. The existence of these groupings continues to suggest that the rate of evolution (albeit on a microevolutionary scale) occurring between transmission seasons and as WNV spreads from region to region is rapid enough to drive the emergence of genetic variants. It is likely to continue to do so as the virus distribution

increases. Although the overall tree topology was conserved between the prM and E protein gene analysis and the complete genome analysis, the phylogenetic relationships of North American West Nile viruses became less clear when using alignments of individual genes or untranslated regions. While analyses of certain genes or untranslated regions were able to generate the topologies illustrated by complete genome or prM and E protein gene analyses (prM, E, NS2A, NS3, NS4B, NS5, and 3'UTR), many did not (5'UTR, capsid, NS1, NS2B, and NS4A). Also, in those analyses that did result in tree topology demonstrating the separation of clades, the resolution of subclades was far less pronounced (Fig. 4-3). The differences in tree topology based on individual genes and untranslated regions may also reflect the ability of some genes or untranslated regions to withstand mutations, while others are unable to withstand mutations because of structural or functional constraints. Alternatively, the rates of evolution of individual genes or untranslated regions may differ because of selective pressures that drive the evolution of certain genes/regions. Thus, while certain individual genes or untranslated regions may be able to act as surrogates to complete genomes when conducting phylogenetic analyses, combinations of genes appear to provide a more accurate depiction of genetic relatedness among isolates in lieu of complete genome sequencing. Despite this finding, it is possible that as more sequence data become available for isolates that the topologies derived from the analysis of each gene or untranslated region will become more consistent. As additional sequence information becomes available from geographic regions underrepresented by this study and in regions where WNV isolates have yet to be obtained for sequencing (i.e., Caribbean islands and Central America), the molecular

epidemiology of this emerging virus will continue to be defined. This dissertation and previous studies (Lanciotti et al., 1999, 2002; Anderson et al., 2001; Ebel et al., 2001, 2004; Huang et al., 2002; Beasley et al., 2003; Davis et al., 2003) have provided a database of sequence information that can be utilized in the future to trace the spread of WNV as its distribution expands. Furthermore, such studies will help to define the modes of transmission and spread that are important for the maintenance and evolution of this virus in nature.

Table 4-1. Isolates used for North American phylogenetic analysis.

Abbreviation	Isolate	Year of isolation	Location	Source	Passage History	Genbank Accession No.
FL 2001 Palm Beach	Florida 2001	2001	Palm Beach, Florida	Catbird	V2	DQ080072
NY 2001 Suffolk	32010157	2001	Suffolk Co., NY	*Cx. pipiens/restuans*	V1	DQ164194
FL 2002 Sumter	Florida 2002	2002	Sumter, Florida	Horse	V2	DQ080071
NY 2002 Queens	02003011	2002	Queens Co., NY	American crow	P	DQ164186
NY 2002 Broome	02003557	2002	Broome Co., NY	American crow	P	DQ164187
NY 2002 Clinton	02002758	2002	Clinton Co., NY	American crow	P	DQ164193
NY 2002 Nassau	34020055	2002	Nassau Co., NY	*Cx. pipiens/restuans*	V2	DQ164195
Ontario 2002	Ontario	2002	Ontario, Canada	American crow	V2	DQ158245
Saskatchewan 2002	Saskatchewan	2002	Saskatchewan, Canada	Black-billed magpie	V2	DQ158246
Manitoba 2002	Manitoba	2002	Manitoba, Canada	Amercican crow	V2	DQ158244
GA 2002 1	68955	2002	Georgia	Human - plasma	P	DQ164196
GA 2002 2	68960	2002	Georgia	Human - brain	P	DQ164197
OH 2002	81948	2002	Ohio	Human - plasma	P	DQ164202
IN 2002	81931	2002	Indiana	Human - plasma	P	DQ164200
TX 2002 1	80025	2002	Texas	Human - plasma	P	DQ164198
TX 2002 2	80022	2002	Texas	Human - plasma	P	DQ164205
NY 2003 Westchester	03000360	2003	Westchester Co., NY	Amercican crow	P	DQ164188
NY 2003 Alabany	03001986	2003	Albany Co., NY	American crow	P	DQ164189
NY 2003 Suffolk	03002018	2003	Suffolk Co., NY	American crow	P	DQ164190
NY 2003 Chautauqua	03002086	2003	Chautauqua Co., NY	American crow	P	DQ164191
NY 2003 Rockland	03002094	2003	Rockland Co., NY	American crow	P	DQ164192
AZ 2003 1	03-az-mp-1623	2003	Arizona	Magpie	V1	DQ158207
AZ 2003 2	03-az-mp-1681	2003	Arizona	Magpie	V1	DQ158208
AZ 2003 3	03-az-mp-1799	2003	Arizona	Magpie	V1	DQ158209
CA 2003 Los Angeles 1	03-ca-crow-s0331532	2003	Los Angeles Co., CA	American crow	V1	DQ158210
CA 2003 Los Angeles 2	03-ca-crow-s0334814	2003	Los Angeles Co., CA	American crow	V1	DQ158211
CA 2003 Los Angeles 3	Mosq. grla1131	2003	Los Angeles Co., CA	*Culex tarsalis*	V2	DQ158212
CA 2003 Los Angeles 4	03-ca-mp-grla-1260	2003	Los Angeles Co., CA	Magpie	V1	DQ158214
CA 2003 Los Angeles 5	Bird 9173	2003	Los Angeles Co., CA	American crow	V2	DQ158217
CA 2003 San Bernadino	Bird 9172	2003	San Bernadino Co., CA	American crow	V2	DQ158218
CA 2003 1	03-ca-mp-impr-102	2003	Imperial Valley, CA	Magpie	V1	DQ158215
CA 2003 2	03-ca-mp-impr-1075	2003	Imperial Valley, CA	Magpie	V1	DQ158216
CA 2003 3	Mosq. impr1143	2003	Imperial Valley, CA	*Culex tarsalis*	V2	DQ158213
MX 2003 Nuevo Leon	Mosq. 9488	2003	Nuevo Leon, Mexico	*Culex quinquefasciatus*	V2	AY963775

Isolates in italics were sequenced by the author.

Table 4-1 (continued). Isolates used for North American phylogenetic analysis.

Abbreviation	Isolate	Year of isolation	Location	Source	Passage History	Genbank Accession No.
CO 2003 Larimer	Bird 9185	2003	Larimer Co., CO	*Culex pipiens*	V2	DQ158219
CO 2003 1	Colorado 3068	2003	Colorado	Red-tailed hawk	P	DQ164204
CO 2003 2	Colorado 3258	2003	Colorado	Magpie	P	DQ164203
NB 2003 Brown	Bird 9239	2003	Brown Co., Nebraska	Bluejay	V2	DQ158236
NB 2003 Dakota	Bird 9241	2003	Dakota Co., Nebraska	Bluejay	V2	DQ158237
TX 2003 Harris 1	Bird 9045	2003	Harris Co., TX	Bluejay	V1	DQ158223
TX 2003 Harris 2	Bird 9114	2003	Harris Co., TX	Bluejay	V1	DQ158222
TX 2003 Harris 3	Mosq. V4095	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	DQ158224
TX 2003 Harris 4	Mosq. V4369	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	AY712948
TX 2003 Harris 5	Mosq. V4096	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	DQ158226
TX 2003 Harris 6	Bird 1153	2003	Harris Co., TX	Mourning dove	V1	AY712945
TX 2003 Harris 7	Bird 1171	2003	Harris Co., TX	Great-tailed grackle	V1	AY712946
TX 2003 Harris 8	Bird 1175	2003	Harris Co., TX	Bluejay	V1	DQ158220
TX 2003 Harris 9	Bird 1240	2003	Harris Co., TX	Bluejay	V1	DQ158221
TX 2003 Harris 10	Bird 1461	2003	Harris Co., TX	Bluejay	V1	AY712947
TX 2003 Harris 11	Mosq. V4370	2003	Harris Co., TX	*Culex quinquefasciatus*	V1	DQ158225
TX 2003 Montgomery 1	Bird 1519	2003	Montgomery Co., TX	Bluejay	V1	DQ158227
TX 2003 Montgomery 2	Bird 1574	2003	Montgomery Co., TX	Bluejay	V1	DQ158228
TX 2003 Montgomery 3	Bird 1576	2003	Montgomery Co., TX	Bluejay	V1	DQ158229
TX 2003 Jefferson 1	Bird 1881	2003	Jefferson Co., TX	Mourning Dove	V1	DQ158230
TX 2003 Jefferson 2	Bird 2073	2003	Jefferson Co., TX	Bluejay	V1	DQ158231
TX 2003 Wharton	Bird 2071	2003	Wharton Co., TX	Bluejay	V1	DQ158232
TX 2003	Texas 82229	2003	Texas	Human - plasma	P	DQ164199
LA 2004 New Iberia	Bird 2409	2004	New Iberia Parish, Louisiana	Northern cardinal	V2	DQ080061
IL 2004 Madison	Illinois - 9515	2004	Madison Co., Illinois	American crow	V2	DQ158250
IL 2004 Schuyler	Illinois - 9517	2004	Schuyler Co., Illinois	American crow	V2	DQ158249
IL 2004 Knox	Illinois - 9519	2004	Knox Co., Illinois	Bluejay	V2	DQ158248
IL 2004 New Marion	Illinois - 9520	2004	New Marion Co., Illinois	Bluejay	V2	DQ158247
AZ 2004	Arizona 2004	2004	Arizona	Human-plasma	P	DQ164201

Table 4-1 (continued). Isolates used for North American phylogenetic analysis.

Abbreviation	Isolate	Year of isolation	Location	Source	Passage History	Genbank Accession No.
MX 2004 Sonora	Mexico 2004	2004	Sonora, Mexico	Human-plasma	V2	AY963774
TX 2004 Harris 1	Bird 2419	2004	Harris Co., TX	Bluejay	V1	DQ158233
TX 2004 Harris 2	Bird 2541	2004	Harris Co., TX	Mourning dove	V1	DQ158234
TX 2004 Harris 3	Bird 3218	2004	Harris Co., TX	Bluejay	V1	DQ158235
TX 2004 Harris 4	Bird 3588	2004	Harris Co., TX	Bluejay	V1	DQ164206
TX 2004 Parmer	Bird 9460	2004	Parmer Co., TX	House sparrow	V1	DQ158238
TX 2004 Brazoria 1	Bird 9461	2004	Brazoria Co., TX	Bluejay	V1	DQ158239
TX 2004 Brazoria 2	Bird 9473	2004	Brazoria Co., TX	Bluejay	V1	DQ158241
TX 2004 Swisher	Bird 9472	2004	Swisher Co., TX	House sparrow	V1	DQ158240
TX 2004 Randall	Bird 9477	2004	Randall Co., TX	House sparrow	V1	DQ158242
TX 2004 Galveston	Bird 9483	2004	Galveston Co., TX	Bluejay	V1	DQ158243

Isolates in bold have been completely sequenced.

P=RNA extracted from infected tissue

V1=RNA extracted from original isolation in Vero cells

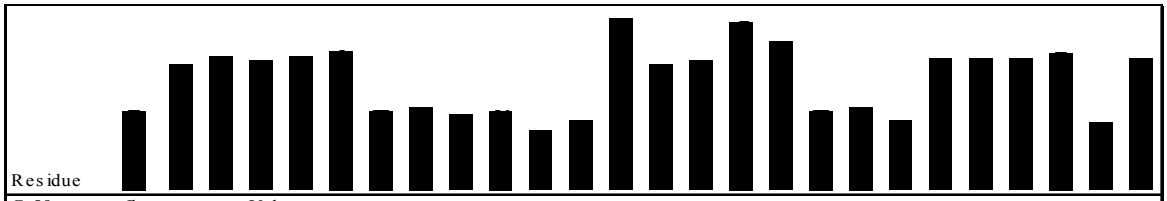
V2=RNA extracted after single Vero cell passage

Isolates in italics were sequenced by author

Table 4-2. Percentage nucleotide/amino acid divergence over the complete genome/ORF among WNV isolates.

	NY99AF196835	NY 2001 Suffolk	NY 2002 Nassau	NY 2002 Clinton	NY 2002 Queens	NY 2002 Broome	GA 2002 1	GA 2002 2	TX 2002 1	TX 2002 2	IN 2002	OH 2002	NY 2003 Westchester	NY 2003 Albany	NY 2003 Suffolk	NY 2003 Chautauqua	NY 2003 Rockland	CO 2003 1	CO 2003 2	TX 2003	TX 2003 Harris 4	TX 2003 Harris 6	TX 2003 Harris 7	TX 2003 Harris 10	AZ 2004	TX 2004 Harris 4
NY99AF196835		**0.15**	**0.18**	**0.18**	**0.17**	**0.24**	**0.24**	**0.23**	**0.23**	**0.23**	**0.16**	**0.30**	**0.18**	**0.25**	**0.28**	**0.22**	**0.26**	**0.32**	**0.34**	**0.25**	**0.24**	**0.33**	**0.35**	**0.25**	**0.39**	**0.21**
NY 2001 Suffolk	0.06		**0.28**	**0.28**	**0.27**	**0.34**	**0.34**	**0.33**	**0.33**	**0.33**	**0.26**	**0.40**	**0.26**	**0.35**	**0.38**	**0.32**	**0.36**	**0.42**	**0.41**	**0.35**	**0.34**	**0.43**	**0.45**	**0.35**	**0.49**	**0.34**
NY 2002 Nassau	0.06	0.12		**0.15**	**0.12**	**0.33**	**0.21**	**0.19**	**0.17**	**0.17**	**0.13**	**0.34**	**0.31**	**0.22**	**0.23**	**0.18**	**0.35**	**0.28**	**0.31**	**0.20**	**0.21**	**0.29**	**0.32**	**0.20**	**0.35**	**0.26**
NY 2002 Clinton	0.06	0.12	0.06		**0.14**	**0.34**	**0.17**	**0.15**	**0.15**	**0.15**	**0.13**	**0.35**	**0.31**	**0.22**	**0.23**	**0.15**	**0.37**	**0.26**	**0.29**	**0.18**	**0.17**	**0.27**	**0.30**	**0.20**	**0.32**	**0.28**
NY 2002 Queens	0.03	0.09	0.03	0.03		**0.32**	**0.20**	**0.18**	**0.16**	**0.16**	**0.10**	**0.33**	**0.30**	**0.21**	**0.20**	**0.17**	**0.34**	**0.27**	**0.30**	**0.19**	**0.20**	**0.28**	**0.31**	**0.19**	**0.34**	**0.25**
NY 2002 Broome	0.15	0.20	0.20	0.20	0.17		**0.41**	**0.39**	**0.37**	**0.37**	**0.33**	**0.44**	**0.36**	**0.42**	**0.43**	**0.38**	**0.19**	**0.48**	**0.49**	**0.40**	**0.41**	**0.45**	**0.48**	**0.40**	**0.55**	**0.41**
GA 2002 1	0.06	0.12	0.06	0.06	0.03	0.20		**0.02**	**0.22**	**0.22**	**0.19**	**0.42**	**0.37**	**0.28**	**0.31**	**0.21**	**0.44**	**0.33**	**0.35**	**0.23**	**0.22**	**0.34**	**0.36**	**0.26**	**0.38**	**0.34**
GA 2002 2	0.03	0.09	0.03	0.03	0.00	0.17	0.03		**0.20**	**0.20**	**0.17**	**0.40**	**0.35**	**0.26**	**0.29**	**0.19**	**0.42**	**0.31**	**0.34**	**0.21**	**0.20**	**0.32**	**0.34**	**0.24**	**0.36**	**0.33**
TX 2002 1	0.09	0.15	0.09	0.09	0.06	0.23	0.09	0.06		**0.02**	**0.17**	**0.38**	**0.35**	**0.26**	**0.27**	**0.19**	**0.40**	**0.29**	**0.32**	**0.21**	**0.22**	**0.30**	**0.33**	**0.23**	**0.36**	**0.31**
TX 2002 2	0.12	0.17	0.12	0.12	0.09	0.26	0.12	0.09	0.03		**0.17**	**0.38**	**0.35**	**0.26**	**0.27**	**0.19**	**0.40**	**0.29**	**0.32**	**0.21**	**0.22**	**0.30**	**0.33**	**0.23**	**0.36**	**0.31**
IN 2002	0.03	0.09	0.03	0.03	0.00	0.17	0.03	0.00	0.06	0.09		**0.34**	**0.29**	**0.20**	**0.23**	**0.16**	**0.35**	**0.26**	**0.31**	**0.20**	**0.19**	**0.27**	**0.30**	**0.20**	**0.34**	**0.26**
OH 2002	0.06	0.12	0.12	0.12	0.09	0.20	0.12	0.09	0.15	0.17	0.09		**0.43**	**0.43**	**0.44**	**0.37**	**0.47**	**0.45**	**0.50**	**0.41**	**0.40**	**0.48**	**0.51**	**0.41**	**0.56**	**0.45**
NY 2003 Westchester	0.09	0.15	0.15	0.15	0.12	0.23	0.15	0.12	0.17	0.20	0.12	0.15		**0.38**	**0.41**	**0.34**	**0.39**	**0.44**	**0.44**	**0.38**	**0.37**	**0.45**	**0.48**	**0.38**	**0.50**	**0.37**
NY 2003 Albany	0.09	0.15	0.09	0.09	0.06	0.23	0.09	0.06	0.12	0.15	0.06	0.15	0.17		**0.32**	**0.25**	**0.44**	**0.35**	**0.40**	**0.29**	**0.28**	**0.36**	**0.39**	**0.29**	**0.43**	**0.35**
NY 2003 Suffolk	0.09	0.15	0.09	0.09	0.06	0.23	0.09	0.06	0.12	0.15	0.06	0.15	0.17	0.12		**0.28**	**0.45**	**0.38**	**0.41**	**0.30**	**0.31**	**0.39**	**0.42**	**0.30**	**0.44**	**0.36**
NY 2003 Chautauqua	0.03	0.09	0.03	0.03	0.00	0.17	0.03	0.00	0.06	0.09	0.00	0.09	0.12	0.06	0.06		**0.41**	**0.26**	**0.33**	**0.22**	**0.19**	**0.31**	**0.34**	**0.24**	**0.35**	**0.32**
NY 2003 Rockland	0.20	0.26	0.26	0.26	0.23	0.12	0.26	0.23	0.29	0.32	0.23	0.26	0.29	0.29	0.29	0.23		**0.51**	**0.52**	**0.43**	**0.44**	**0.50**	**0.53**	**0.43**	**0.58**	**0.44**
CO 2003 1	0.03	0.09	0.03	0.03	0.00	0.17	0.03	0.00	0.06	0.09	0.00	0.09	0.12	0.06	0.06	0.00	0.23		**0.44**	**0.34**	**0.31**	**0.32**	**0.34**	**0.34**	**0.44**	**0.42**
CO 2003 2	0.15	0.20	0.15	0.15	0.12	0.29	0.15	0.12	0.17	0.20	0.12	0.20	0.23	0.17	0.17	0.12	0.35	0.12		**0.34**	**0.35**	**0.45**	**0.48**	**0.36**	**0.44**	**0.43**
TX 2003	0.06	0.12	0.06	0.06	0.03	0.20	0.06	0.03	0.09	0.12	0.03	0.12	0.15	0.09	0.09	0.03	0.26	0.03	0.15		**0.24**	**0.34**	**0.37**	**0.25**	**0.39**	**0.34**
TX 2003 Harris 4	0.12	0.17	0.12	0.12	0.09	0.26	0.12	0.09	0.15	0.17	0.09	0.17	0.20	0.15	0.15	0.09	0.32	0.09	0.20	0.12		**0.34**	**0.36**	**0.26**	**0.38**	**0.34**
TX 2003 Harris 6	0.12	0.17	0.12	0.12	0.09	0.26	0.12	0.09	0.15	0.17	0.09	0.17	0.20	0.15	0.15	0.09	0.32	0.09	0.20	0.12	0.17		**0.03**	**0.34**	**0.48**	**0.43**
TX 2003 Harris 7	0.17	0.23	0.17	0.17	0.15	0.32	0.17	0.15	0.20	0.23	0.15	0.23	0.26	0.20	0.20	0.15	0.38	0.15	0.26	0.17	0.23	0.06		**0.37**	**0.51**	**0.45**
TX 2003 Harris 10	0.15	0.20	0.15	0.15	0.12	0.29	0.15	0.12	0.17	0.20	0.12	0.20	0.23	0.17	0.17	0.12	0.35	0.12	0.23	0.15	0.20	0.20	0.26		**0.41**	**0.32**
AZ 2004	0.17	0.23	0.17	0.17	0.15	0.32	0.17	0.15	0.20	0.23	0.15	0.23	0.26	0.20	0.20	0.15	0.38	0.15	0.20	0.17	0.23	0.23	0.29	0.26		**0.49**
TX 2004 Harris 4	0.09	0.15	0.09	0.09	0.06	0.23	0.09	0.06	0.12	0.15	0.06	0.15	0.17	0.12	0.12	0.06	0.29	0.06	0.17	0.09	0.15	0.15	0.20	0.12	0.20	

Table 4-3. Deduced amino acid substitutions from NY-99 found in the ORF of each completely sequenced isolate.



Residue																												
C-59	Ile			Val																								
C-67	Asp	Asn																										
C-90	Thr																										Ala	
prM-4	Asn																					Asp						
prM-156	Val																						Ile	Ile				
E-159	Val		Ala	Ala	Ala		Ala	Ala	Ala	Ala	Ala			Ala	Ala	Ala		Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala	Ala
E-169	Ile																										Val	
E-176	Tyr					His											His											
E-178	Leu											Gln																
E-188	Thr																											Ala
E-208	Thr																											
E-263	His																Gln											
NS1-111	Leu																Phe											
NS1-135	Val					Ile											Ile											
NS1-144	Pro	Ser																										
NS1-254	His																		Tyr									
NS2A-34	Met					Leu											Leu											
NS2A-102	Gln								His	His																		
NS2A-224	Ala											Thr																
NS2B-84	Asn										Asp																	
NS3-106	Val																Ala											
NS3-328	Glu																									Lys		
NS3-336	Leu																											
NS3-486	Phe												Leu														Ser	
NS3-539	Arg						Lys																					
NS3-603	Ser																											
NS4A-3	Ile		Met									Leu																
NS4A-85	Ala																		Thr								Thr	
NS4A-135	Val																									Met		
NS4B-83	Ala										Ser																	
NS4B-119	Phe					Leu											Leu											
NS4B-202	Ile																		Thr									
NS4B-240	Ile								Met	Met																		
NS4B-241	Thr																					Ala						
NS4B-245	Ile																		Val									
NS4B-249	Glu																						Gly	Gly				
NS5-200	Arg																							Leu				
NS5-277	Ser														Arg													
NS5-287	Arg	Lys																										
NS5-296	His																					Tyr						
NS5-314	Lys																										Arg	
NS5-531	Lys												Arg															
NS5-560	Ala																											Ser
NS5-619	Ala																									Ser		
NS5-688	Ala																							Asp				
NS5-804	Ala																						Val	Val				
Total		3	2	2	1	4	2	1	3	3	1	2	3	3	2	1	7	1	5	1	4	4	6	4	6	3		

Amino acid positions determind from the sequence alignment of the 25 completely sequenced isolates in comparison to WN-NY99

Table 4-4. Nucleotide mutations conserved in all isolates of North American clade 2002-2004 relative to WN-NY99.

5'UTR	Capsid	prM	E	NS1	NS2A	NS2B	NS3	NS4A	NS4B	NS5	3'UTR
None	None	660 (C to U)	1442 (U to C)*	None	3774 (U to C)	None	4803 (C to U)	None	6996 (C to U)	7938 (U to C)	10851
			2466 (C to U)		4146 (A to G)		6138 (C to U)			9352 (C to U)	(A to G)
							6238 (C to U)				
							6426 (C to U)				

Nucleotide positions correspond to WN-NY99.
*Encodes E159 (Val to Ala) amino acid substitution.

Figure 4-1. Phylogenetic tree constructed by Bayesian analysis of prM and E protein genes (2004 nucleotides) of 108 North American isolates. Numbers at each node represent bootstrap support following 500 replicates.

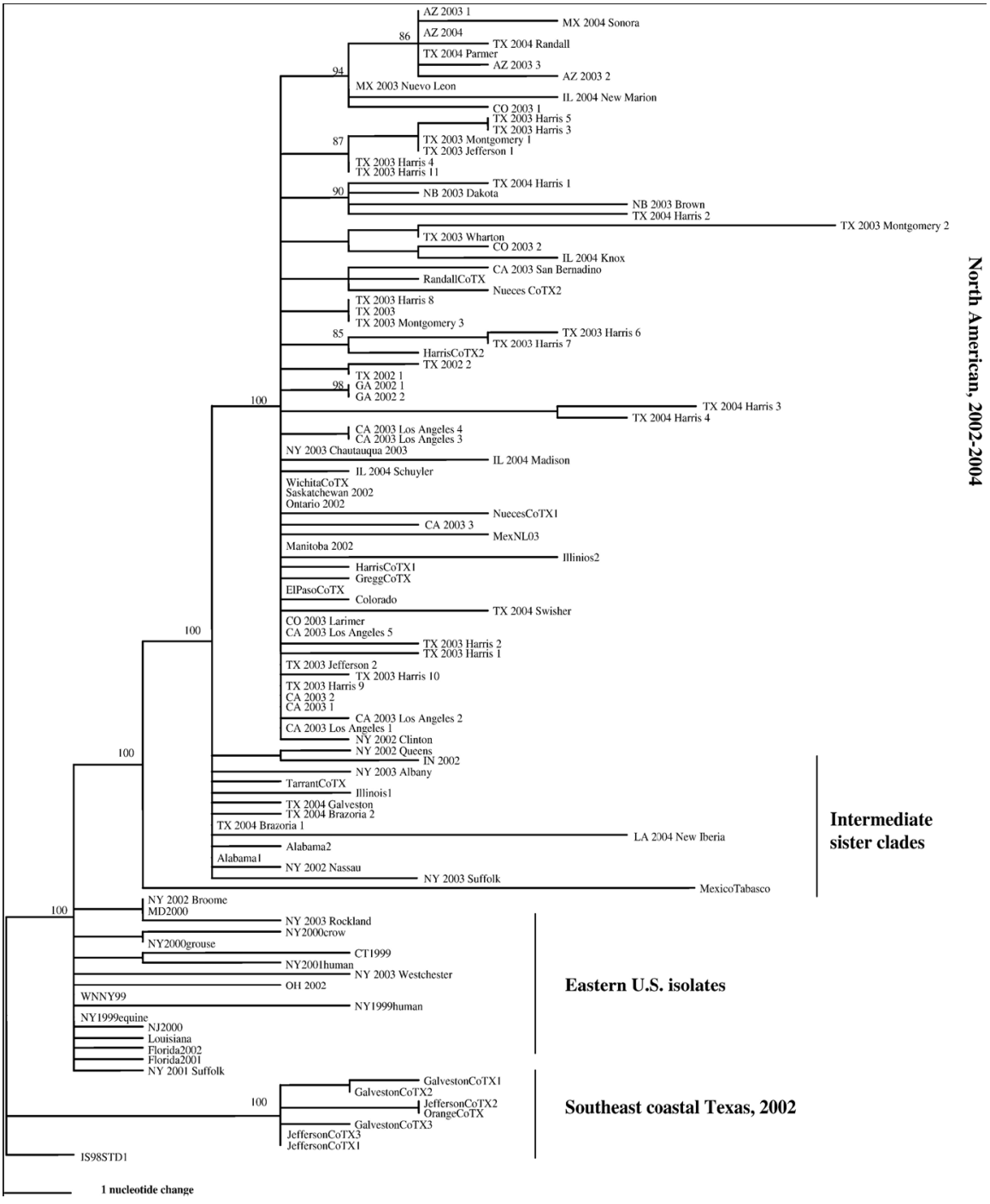


Figure 4-2. Phylogenetic tree generated by maximum likelihood analysis of an alignment of the complete genomes of 25 newly sequenced North American WNV isolates with previously sequenced U.S. and Old World isolates. Numbers at each node represent bootstrap support following 500 replicates.

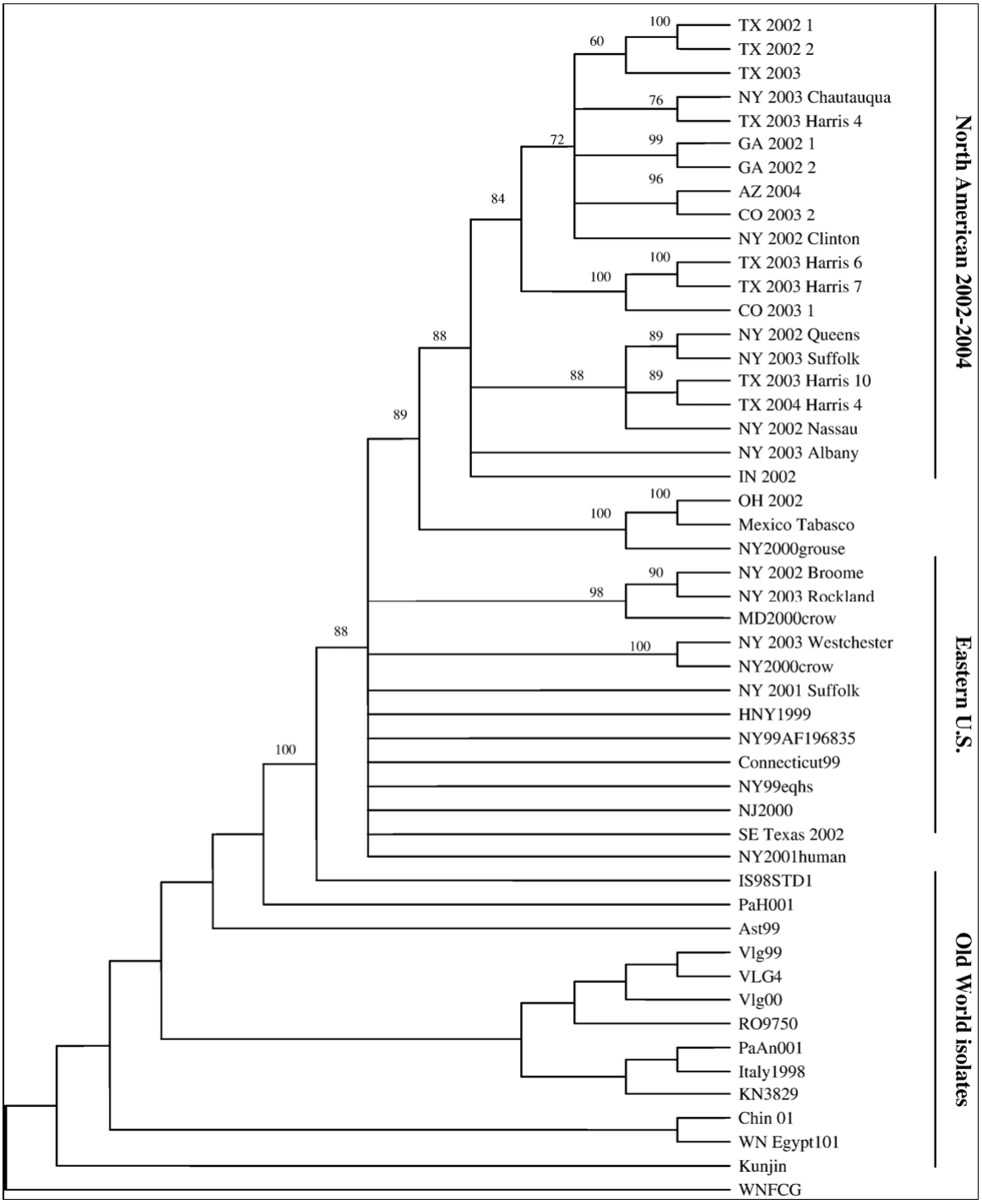
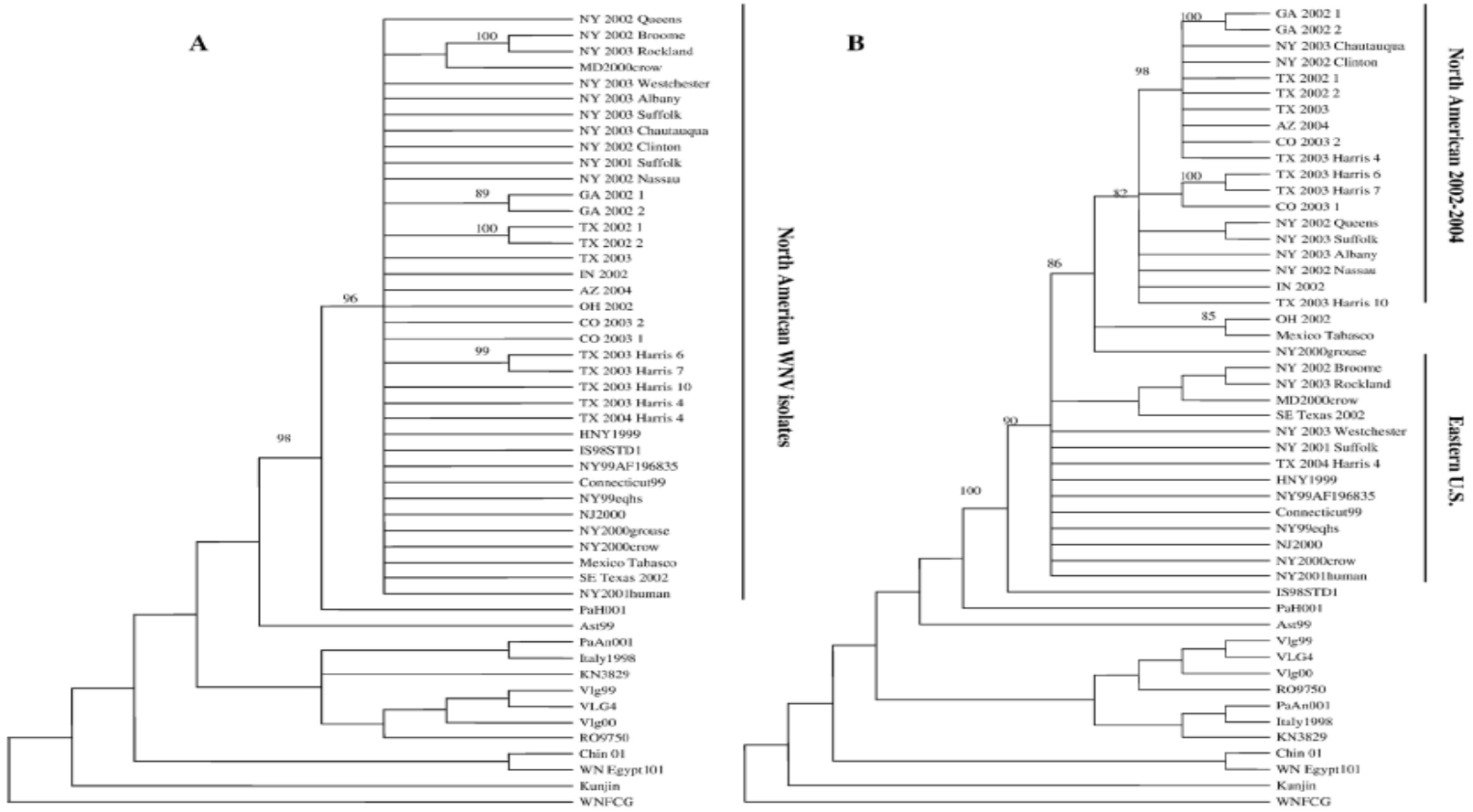


Figure 4-3. Representative phylogenetic trees generated by maximum likelihood analysis of each of the individual WNV genes and untranslated regions showing presence or absence of clade topology with North American, 2002-2004 clade separate from Eastern U.S. clade.

Group "A" represents tree topology generated from analyses of 5'UTR, capsid, NS1, NS2B, and NS4A. Group "B" represents tree topology generated from analyses of prM, E, NS2A, NS3, NS4B, NS5, and 3'UTR.



CHAPTER 5

IDENTIFICATION OF PHENOTYPIC VARIATION IN NORTH AMERICAN WEST NILE VIRUS VARIANTS

Introduction

Because of the relatively recent emergence of West Nile virus in North America and its continued spread, studies concerning the evolution of WNV are important to understand the extent to which the virus has mutated as its temporal and geographic distribution have expanded. Following the introduction of WNV into Texas during 2002, studies were initiated to determine if phenotypic changes also occurred among the previously described genetic variants. Although genetically distinct on a microevolutionary scale, previous studies indicated that the phenotypic characteristics (i.e., plaque morphology, *in vitro* growth kinetics, neuroinvasiveness and neurovirulence in a mouse model) of WNV isolates collected in 2002 were not significantly different from isolates collected in 1999 (Beasley et al., 2003). During the 2003 transmission season, however, isolates of WNV from both birds and mosquitoes collected in Texas were recovered by Dr. Robert Tesh and Ms. Hilda Guzman that produced small plaque (sp) morphology and reduced virus yield at 72 hours following infection in Vero cells in comparison to isolates of WNV from previous transmission seasons. These and other

isolates were characterized to define the first description of phenotypic variation among WNV isolates in North America.

Results

Identification of small plaque and temperature sensitive isolates

Six of 29 isolates made during the 2003 transmission season were identified with a small plaque (sp) phenotype (Figure 5-1). Specifically, WN-NY99 and other WNV isolates had a large plaque morphology (>1.5 mm), while a sp morphology was identified in 2003 (< 1.0 mm in diameter). All sp isolates were collected in Harris Co. or Montgomery Co., TX and were made over a four-month period (May 09-Sept. 8, 2003). Each isolate was given a single passage in Vero cells following initial isolation. Previous studies with other related flaviviruses (yellow fever, dengue-2, tick-borne encephalitis, St. Louis encephalitis, and Japanese encephalitis viruses) have shown that some viruses with sp phenotypes are also temperature-sensitive (ts) when grown in cell culture (Blaney et al., 2003a; Ledger et al., 1992; Hollingshead et al., 1983; Eastman and Blair, 1985; Wallner et al., 1996). Thus, experiments were conducted to determine if sp and non-sp producing isolates of WNV collected in 1999, 2002, and 2003 displayed a ts phenotype (Table 5-1). Three isolates (Bird 1153, Bird 1171, and Bird 1175) exhibiting a sp phenotype were identified as also having a ts phenotype at 39.5°C compared to the permissive growth temperature of 37.0°C, while the other 26 sp and non-sp WNV

isolates collected during 1999, 2002, and 2003 did not. Temperature-sensitivity assays were also carried out on sp isolates at 41.0°C in order to test if isolates not sensitive to 39.5°C were sensitive to higher temperatures (Table 5-1). Although two of the three sp isolates sensitive to 39.5°C did reveal a greater degree of sensitivity to 41.0°C (approximately 0.2-0.3 $\log_{10}$ decrease in plaque titer at 41.0°C), all other non-ts isolates remained non-ts at 41.0°C.

Multiplication kinetics in cell culture

Following the identification of isolates with sp and/or ts phenotypes, *in vitro* multiplication studies in Vero cells at an MOI of 0.1 pfu/cell were undertaken to compare these isolates to WN-NY99 (strain 382-99) and a 2002 large plaque (lp), non-ts Texas WNV isolate (Bird 113). Results from this analysis suggested that viral isolates with sp and/or ts phenotypes have reduced levels of replication in Vero cells at early time points in comparison to WN-NY99 and isolates from 2002 with lp phenotypes (Figure 5-2). These isolates also exhibited a 10 to 100-fold decrease in infectivity titer at each time point during the course of infection and did not reach the peak viral titers observed for other WNV isolates at 72 hours postinfection.

Mouse attenuation phenotypes

The identification of viral isolates exhibiting sp and/or ts phenotypes with reduced replication in cell culture led to the investigation of the virulence phenotypes of these isolates, using a mouse model. Small plaque and ts phenotypes have previously been shown to be indicative of attenuation of neuroinvasiveness and neurovirulence in mouse models for many of the encephalitic flaviviruses (Hanley et al., 2003; Hanley et al., 2002; Blaney et al., 2003; Puri et al., 1997; Wallner et al., 1996). Consequently, each of the 2003 WNV isolates that were tested for sp and ts phenotypes were also tested for attenuation of neuroinvasiveness and neurovirulence in female 3-4 week-old Swiss Webster mice. Using this model, it has been possible to determine both intraperitoneal (i.e., neuroinvasiveness) and intracerebral (i.e., neurovirulence) virulence values of several U.S. WNV isolates collected prior to 2003. Because the intraperitoneal (ip) LD_{50} of U.S. WNV isolates from 1999 and 2002 ranged from 0.4-4.2 pfu (Beasley et al., 2003), WNV isolates collected in 2003 were screened for attenuation of neuroinvasiveness by administering intraperitoneal doses of either 10^2 or 10^3 pfu of each isolate (Table 5-2). Six isolates were identified as being attenuated for neuroinvasiveness (ip $LD_{50} \geq 1{,}000$ pfu) (Bird 1171, Bird 1175, Bird 1153, Bird 1479, Bird 1519, and Mosq. v4369) the extent of attenuation of neurovirulence was examined for three of the attenuated isolates (Bird 1153, Bird 1519, Mosq. v4369) by intracerebral (ic) inoculation. The results of these studies revealed that these three isolates were attenuated for mouse neuroinvasiveness but were not attenuated for mouse neurovirulence in this model (Table 5-2).

Attenuation of mouse neuroinvasiveness

In order to establish the degree of attenuation of neuroinvasiveness in a Swiss Webster mouse model, precise neuroinvasiveness for sp and/or ts isolates that exhibited a ≥1,000 pfu per LD_{50} was determined. The extent of attenuation of neuroinvasiveness ranged from an ip LD_{50} of 2,000 pfu (Bird 1171) to 645,000 pfu (Mosq. v4369) representing up to 100,000-fold attenuation when compared to the prototypic WNV isolate from New York in 1999 (Table 5-3). Also, the percent mortality following infection with each isolate was greater at higher doses, suggesting the dose-dependent manifestation of encephalitis in this mouse model. At 21 days post-inoculation (dpi), surviving mice were challenged with 100 LD_{50} of WN-NY99 (strain 385-99) to measure the dose of each isolate required to protect 50% of mice from death following challenge with a highly neuroinvasive strain (PD_{50}). These experiments suggest that attenuated isolates administered at even very low doses (10^{-1} to 10^{0} pfu) are able to induce a protective immune response in the mouse and that although attenuated, these isolates must undergo some replication (Table 5-3). To investigate the mechanism by which these isolates were attenuated, serum viremia curves and brain infectivity levels were determined for a single sp, ts, mouse attenuated WNV isolate (Bird 1153) and compared to the levels of viremia in the mouse following parallel infection with WN-NY99. Groups of mice were infected ip with 10^{3} pfu of either WN-NY99 (strain 382-99) or isolate Bird 1153, and three mice from each group were sacrificed daily for 8 days for collection of serum and whole brain preparations. As expected, WN-NY99 followed the typical course

of infection of a highly neuroinvasive WNV, whereby serum viremia increased daily following inoculation, peaked at day three, and was followed by the detection of virus in the brain as early as day four and continued until death at day seven-eight (Table 5-4). Isolate Bird 1153, however, did not produce increased levels of serum viremia following infection. Viral titers remained at levels comparable to the titer of the inoculum for at least 24 hours post-inoculation, but decreased by 48 or 72 hours and became undetectable in the serum by day four. At an ip dose of 10^3 pfu, no virus was detectable in the brains of these mice at any time post inoculation. These results suggest that attenuation of at least one of the non-neuroinvasive 2003 WNV isolates is due to the inability of the virus to replicate to high levels outside the brain sufficient for the invasion of the central nervous system of the mouse.

Phenotypic characterization of 2004 West Nile virus isolates

A total of 34 isolates made from several locations in North America during 2004 were characterized for the sp/ts phenotype. All isolates obtained were lp and non-ts at 41.0ºC (Table 5-5). Because of the high correlation between the sp and ts phenotypes and mouse attenuation, none of the lp and non-ts isolates were tested for attenuation in the mouse model.

Discussion

These studies provide the first evidence of phenotypic variation in the North American WNV population by the characterization of isolates exhibiting small plaque morphology, temperature-sensitivity, and attenuation in a mouse model. Previous mouse virulence studies have shown that both lineage 1 and 2 WNV isolates made from nature (i.e., mosquito pools, birds, horses, humans) differ in their ability to replicate in peripheral tissues, induce viremia, and invade the CNS (neuroinvasiveness), but that all strains are able to initiate a cytopathic infection in the CNS and cause encephalitis (neurovirulence) if virus is delivered directly to the brain via ic inoculation (Beasley et al., 2002). The present studies have shown similar findings in that naturally acquired isolates, which were attenuated for mouse neuroinvasiveness, were not attenuated for mouse neurovirulence in the current model. Additionally, the manifestation of encephalitic disease in this mouse model following infection with the 2003 attenuated isolates appeared to be dose-dependent with only the highest doses ($\geq 10^3$ pfu per dose) causing disease, suggesting that invasion of the CNS is determined by the amount of virus (level of viremia) circulating in the mouse. While this observation may be relevant for this mouse model, other species, such as birds and hamsters, are able to support higher viremias, yet remain asymptomatic, while humans and horses may develop encephalitis with a relatively low transient viremia (Bunning et al., 2002; Komar et al., 2003; Xiao et al., 2001). Thus, it is likely that there are multiple virulence determinants in different species that may influence the phenotypes of these isolates in various animal models.

Although the experiments in this chapter suggest that attenuation of these isolates is related to their inability to replicate to titers sufficient for the invasion of the CNS, the underlying mechanism of attenuation remains unclear. Isolates from 2003 had a range of phenotypic properties (Table 5-1, Table 5-2). Therefore, genetic analysis of multiple variants is likely to identify multiple mutations that affect a range of phenotypic properties, including mouse virulence phenotypes. For those isolates that exhibited a ts phenotype, it is possible that the mechanism of attenuation was the result of viral replication being sensitive to the temperature of the mouse following infection and induction of fever. Although ts variants of WNV and other flaviviruses have been produced *in vitro* by either cell culture passage or chemical mutagenesis (Blaney et al., 2003b; Dunster et al., 1990; Hanley et al., 2002), to the knowledge of the author, the natural occurrence of WNV isolates exhibiting temperature sensitivity has not been previously reported. The higher body temperatures of avian species may be a selective factor in the emergence of ts variants. Additional *in vivo* bird experiments will be necessary to address this question. It has been hypothesized that mutations to viral genes that encode proteins with enzymatic functions may result in generation of virus variants with a temperature-sensitive phenotype. This is because every enzyme has an optimal temperature for activity, leading to a decline in its functional role in viral replication at suboptimal temperatures. Similarly, mutations may destabilize the conformation of the enzyme by breaking bonds between amino acids important in the active site and protein stability. Conformational instability of the enzyme may then result in suboptimal enzyme-substrate binding. In support of this hypothesis, mutations in the nonstructural

protein genes (i.e., NS1, NS3, NS5) of mosquito-borne flaviviruses have been correlated with the generation of a temperature-sensitive phenotype (Blaney et al., 2003; Hanley et al., 2002; Muylaert et al., 1997). However, it is possible that mutations in structural protein genes may also result in a temperature-sensitive phenotype if the conformational stability of, for example, the envelope protein is affected. Experiments in the following chapter of this dissertation will utilize complete genome sequencing and reverse genetic studies to identify the molecular determinants of the phenotypes displayed by these isolates. Finally, it is unknown whether or not the attenuated phenotype of these isolates will be retained in other susceptible hosts, such as birds, horses, or humans. However, it would be expected that variants will arise that are attenuated for a variety of hosts, including birds, equines, and/or humans. Further studies are warranted to investigate if there is a selection for WNV variants with modified phenotypes and should stress the importance of continued isolation of WN viruses from the field as both a surveillance tool and a means of studying the evolution of this virus over time. Continued genetic and phenotypic characterization of WNV isolates made in North America will improve our understanding of the relationship between viral evolution and phenotypic variation in an emerging viral population and how these factors influence the epidemiology of WNV in North America.

Table 5-1. Small plaque and temperature-sensitivity phenotypes of West Nile viruses isolated in Texas, 2003.

	Virus	Source	Small plaque	Mean virus titer (log_{10} PFU/ml) at indicated temperature (°C) Vero				
				37°C	39.5°C	Δ	41.0°C	Δ
	WN-NY99 (382-99)	Flamingo brain	-	8.1	7.6	0.5	8.3	0.2
	WN-99 (385-99)	Owl brain	-	7.0	7.3	0.3	7.2	0.2
	TWN 93 (2002)	Bird 113	-	7.9	7.7	0.2	8.0	0.1
	TWN 84 (2002)	Bird 114	-	7.4	7.6	0.2	7.8	0.4
	TWN 117 (2002)	Bird 476 (Bolivar)	-	7.0	7.3	0.3	7.3	0.3
2003 WNV Isolates	TWN 269	Bird 1171	***sp***	6.0	3.3	**2.7**	3.0	**3.0**
	TWN 270	Bird 1175	***sp***	3.5	1.0	**2.5**	1.0	**2.5**
	TWN 271	Bird 1240	-	3.5	3.4	0.1	3.6	0.1
	TWN 274	Bird 1153	***sp***	7.7	5.0	**2.7**	4.8	**2.9**
	TWN 300	Bird 1427	-	5.4	5.9	0.5	5.4	0.0
	TWN 301	Bird 1461	-	7.0	6.7	0.3	7.2	0.2
	TWN 302	Bird 1479	***sp***	6.3	7.0	0.7	7.0	0.7
	TWN 303	Bird 1481	-	6.0	6.2	0.2	6.0	0.0
	TWN 304	Bird 1496	-	6.2	6.7	0.5	6.4	0.2
	TWN 305	Bird 1519	***sp***	6.3	6.6	0.3	6.5	0.2
	TWN 358	Mosq. v2769	-	6.1	6.3	0.2	6.5	0.4
	TWN 359	Mosq. v3437	-	6.2	6.6	0.4	6.4	0.2
	TWN 360	Mosq. v3567	-	7.1	6.7	0.4	7.0	0.1
	TWN 361	Mosq. v3693	-	7.2	7.2	0.0	7.4	0.2
	TWN 362	Mosq. v3941	-	7.2	7.0	0.2	7.4	0.2
	TWN 363	Mosq. v3942	-	7.2	7.3	0.1	7.6	0.4
	TWN 364	Mosq. v4007	-	8.0	8.0	0.0	8.2	0.2
	TWN 365	Mosq. v4095	-	7.5	7.7	0.2	7.3	0.2
	TWN 366	Mosq. v4181	-	7.9	8.0	0.1	7.9	0.0
	TWN 367	Mosq. v4195	-	7.7	7.9	0.2	7.9	0.2
	TWN 378	Bird 2071	-	8.0	7.9	0.1	8.3	0.3
	TWN 379	Bird 2073	-	7.4	7.6	0.2	7.6	0.2
	TWN 382	Mosq. v4369	***sp***	7.1	5.3	1.8	5.5	1.6
	TWN 383	Mosq. v4380	***mixed***	6.6	6.3	0.3	6.5	0.1
	TWN 399	Bird 2066	-	7.8	7.7	0.1	7.5	0.3
	TWN 400	Bird 2067	-	7.5	7.1	0.4	7.3	0.2
	TWN 401	Bird 2075	-	7.9	7.7	0.2	7.8	0.1
	TWN 402	Bird 1556	-	7.5	7.3	0.2	7.5	0.0
	TWN 404	Bird 1881	***mixed***	6.4	7.0	0.3	6.9	0.5

sp = small plaque size at 37°C; small plaques have a diameter of < 1.0mm compared to diameter of >1.5mm
mixed = both small and large plaques measured
Δ = Change in titer (log_{10}PFU/ml) at 39.5°C or 41.0°C compared to titer at permissive temperature (37.0°C).
Underline indicates >2.0 log_{10} PFU/ml reduction in titer at 39.5°C or 41.0°C when compared to 37.0°C.

Table 5-2. Mouse attenuation phenotypes of West Nile viruses isolated in Texas, 2003

	Virus	Source	Mouse neuroinvasiveness and neurovirulence: Intraperitoneal inoculation LD_{50} (PFU)	Intraperitoneal inoculation AST ± s.d.(days)	Intracerebral inoculation LD_{50} (PFU)	Intracerebral inoculation AST ± s.d.(days)
	WN-NY99 (382-99)	Flamingo brain	0.8	8.0 ± 1.2	**0.1**	**6.4 ± 0.9**
	WN-99 (385-99)	Owl brain	0.4	7.2 ± 0.6	**0.4**	**6.0 ± 0.2**
	TWN 93 (2002)	Bird 113	0.5	8.0 ± 1.0	**0.1**	**6.7 ± 0.3**
	TWN 84 (2002)	Bird 114	4.2	7.5 ± 1.2	n.d.	n.d.
	TWN 117 (2002)	Bird 476 (Bolivar)	1.8	8.6 ± 0.7	n.d.	n.d.
2003 WNV Isolates	TWN 269	Bird 1171	**≥1,000**	n/a	n.d.	n.d.
	TWN 270	Bird 1175	**≥1,000**	n/a	n.d.	n.d.
	TWN 271	Bird 1240	≤100	8.0 ± 2.0	n.d.	n.d.
	TWN 274	Bird 1153	**≥1,000**	n/a	**0.3**	**6.2 ± 0.4**
	TWN 300	Bird 1427	≤100	7.0 ± 1.0	n.d.	n.d.
	TWN 301	Bird 1461	≤100	7.0 ± 1.0	**0.2**	**6.2 ± 0.6**
	TWN 302	Bird 1479	**≥100**	n/a	n.d.	n.d.
	TWN 303	Bird 1481	≤100	7	n.d.	n.d.
	TWN 304	Bird 1496	≤100	7.2 ± 1.0	n.d.	n.d.
	TWN 305	Bird 1519	**≥1,000**	n/a	**0.1**	**6.9 ± 0.2**
	TWN 358	Mosq. v2769	≤100	7.5 ± 3.0	n.d.	n.d.
	TWN 359	Mosq. v3437	≤100	7.0 ± 1.0	n.d.	n.d.
	TWN 360	Mosq. v3567	≤100	8.6 ± 2.0	n.d.	n.d.
	TWN 361	Mosq. v3693	≤100	8.4 ± 1.0	n.d.	n.d.
	TWN 362	Mosq. v3941	≤100	7.3 ± 1.0	n.d.	n.d.
	TWN 363	Mosq. v3942	≤100	8.2 ± 1.0	n.d.	n.d.
	TWN 364	Mosq. v4007	≤100	8.0 ± 2.0	n.d.	n.d.
	TWN 365	Mosq. v4095	≤100	8.4 ± 1.5	n.d.	n.d.
	TWN 366	Mosq. v4181	≤100	11 ± 3.7	n.d.	n.d.
	TWN 367	Mosq. v4195	≤100	7.2 ± 1.3	n.d.	n.d.
	TWN 378	Bird 2071	≤100	8.4 ± 1.2	n.d.	n.d.
	TWN 379	Bird 2073	≤100	8.1 ± 2.0	n.d.	n.d.
	TWN 382	Mosq. v4369	**≥1,000**	n/a	**0.1**	**5.8 ± 0.4**
	TWN 383	Mosq. v4380	**≥100**	n/a	n.d.	n.d.
	TWN 399	Bird 2066	≤100	7.2 ± 0.6	n.d.	n.d.
	TWN 400	Bird 2067	≤100	7.1 ± 0.2	n.d.	n.d.
	TWN 401	Bird 2075	≤100	7.0 ± 1.3	n.d.	n.d.
	TWN 402	Bird 1556	≤100	7.8 ± 2.0	n.d.	n.d.
	TWN 404	Bird 1881	**≥1,000**	n/a	n.d.	n.d.

For a dose of 10^3 PFU of virus; only strains causing mortality in at least 4 of 5 animals have average survival time (AST).
≥1,000 indicates attenuated phenotype.
≥100 = LD_{50} >100 PFU but <1,000 PFU.
Note . n.d., not determined

Table 5-3. Mouse neuroinvasive/neurovirulence phenotypes.

West Nile virus isolate	ip LD_{50} (PFU)	ic LD_{50} (PFU)	A.S.T. ±s.d. (*P*)	ip PD_{50} (PFU)
WN-NY99 (382-99)	0.8	0.1	7.5 ± 0.6	1.3
Texas 2002 (Bird 113)	0.5	0.1	8.0 ± 1.0 (0.2)*	1.5
Texas 2003 (Bird 1461)	0.6	0.2	7.0 ± 1.0 (0.7)*	1.2
Texas 2003 (Bird 1153)	23,000	0.3	9.5 ± 1.0 (0.12)*	0.8
Texas 2003 (Bird 1519)	51,000	0.1	9.0 ± 4.0 (0.15)*	0.4
Texas 2003 (Mosquito 4369)	645,000	0.1	8.3 ± 3.0 (0.4)*	0.1
Texas 2003 (Bird 1171)	2,000	n.d.	9.7 ± 3.3 (0.1)*	1.0
Texas 2003 (Bird 1175)	10,000	n.d.	9.3 ± 6.0 (0.13)*	0.2

A.S.T. ± s.d. (ip) = Average survival time.

ip PD_{50} = the number of pfu required to immunize mice by the ip route to protect against challenge with 100 LD_{50} WN-NY99 (382-99).

P values determine by two-tailed Mann-Whitney test compared to WN-NY99. Asterisks indicate no significance.

n.d. = not determined

Table 5-4. Serum and brain viremia in mice following ip inoculation of 10^3 pfu of WN-NY99 (neuroinvasive) vs. Bird 1153 (non-neuroinvasive).

Serum and brain viremia in mice following ip inoculation of 10^3 pfu of WN-NY99 (neuroinvasive) vs. Bird 1153 (non-neuroinvasive)

Days post inoculation	Animal	WN-NY99 (382-99)		Bird 1153	
		Serum titer (pfu/ml)	Brain titer (pfu/brain)	Serum titer (pfu/ml)	Brain titer (pfu/brain)
1	1	2,000	-	3,000	-
	2	4,000	-	1,500	-
	3	1,400	-	4,000	-
2	1	4,000	-	1,650	-
	2	11,500	-	250	-
	3	1,650	-	750	-
3	1	17,000	-	100	-
	2	15,000	-	100	-
	3	6,000	-	250	-
4	1	-	2,000	-	-
	2	-	500	-	-
	3	-	4,000	-	-
5	1	-	3,000	-	-
	2	-	1,000	-	-
	3	-	750	-	-
6	1	-	135,000	-	-
	2	-	300,000	-	-
	3	-	750,000	-	-
7	1	-	1,150,000	n.d.	n.d.
8	1	-	1,350,000	n.d.	n.d.

*** - indicates no virus detected**

Table 5-5. Plaque morphology and temperature-sensitivity of 2004 WNV isolates

	Virus	Source	Small plaque	Mean virus titer (log10 PFU/ml) at indicated temperature (°C) Vero 37°C	Vero 41.0°C	Vero Δ
2004 WNV Isolates	TWN 496	Bird 2409	-	6.3	6.0	0.3
	TWN 501	Bird 2419	-	7.2	7.5	0.3
	TWN 532	Bird 2514	-	7.4	7.5	0.1
	TWN 577	Bird 3218	-	7.6	7.9	0.3
	TWN 578	Bird 3588	-	7.6	7.8	0.2
	TWN 600	Bird 3683	-	7.2	7.4	0.2
	TWN 601	Bird 3678	-	7.7	7.0	0.7
	TWN 602	Bird 3680	-	5.4	5.6	0.2
	TWN 603	Bird 3681	-	6.7	6.9	0.2
	TWN 605	TVP-9477 (Bird)	-	6.3	6.5	0.2
	TWN 616	Bird 3567	-	7.1	7.5	0.4
	TWN 617	Bird 3838	-	6.8	6.9	0.1
	TWN 619	TVP-9498 (Human)	-	7.1	7.0	0.1
	TWN 637	TVP-9515 (Bird)	-	7.2	7.4	0.2
	TWN 638	TVP-9517 (Bird)	-	6.2	6.8	0.6
	TWN 639	TVP-9519 (Bird)	-	7.2	7.6	0.4
	TWN 640	TVP-9520 (Bird)	-	8.0	8.2	0.2
	TWN 675	TVP-9613 (Bird)	-	7.5	7.3	0.2
	TWN 676	TVP-9614 (Bird)	-	7.9	7.9	0.0
	TWN 679	TVP-9609 (Bird)	-	6.8	6.9	0.1
	TWN 680	TVP-9610 (Bird)	-	5.6	5.9	0.3
	TWN 720	TVP-9663 (Bird)	-	5.9	6.3	0.4
	TWN 721	TVP-9656 (Bird)	-	6.2	6.6	0.4
	TWN 722	TVP-9664 (Bird)	-	7.5	7.4	0.1
	TWN 723	TVP-9658 (Bird)	-	7.8	7.3	0.5
	TWN 724	TVP-9677 (Bird)	-	7.5	7.7	0.2
	TWN 725	TVP- 9657 (Bird)	-	6.3	6.4	0.1
	TWN 726	TVP-9660 (Bird)	-	6.7	6.9	0.2
	TWN 727	TVP-9665 (Bird)	-	8.2	8.2	0.0
	TWN 732	TVP-4511 (Bird)	-	7.8	8.0	0.2
	TWN 733	TVP-4486 (Bird)	-	6.8	6.9	0.3
	TWN 734	TVP-4276 (Bird)	-	7.4	7.5	0.1
	TWN 735	TVP-4487 (Bird)	-	7.2	7.1	0.1

sp = small plaque size at 37°C; small plaques have a diameter of < 1.0mm compared to diameter of >1.5mm
Δ = Change in titer (log_{10}PFU/ml) at 41.0°C compared to titer at permissive temperature (37.0°C).

Fig. 5-1. Plaque morphology of WNV isolates. Vero cells in 6-well plates were infected with WN-NY99, WNV 2002, WNV 2003 sp.

Plaques were visualized 3 days postinoculation by staining with crystal violet. Images from pictures copied to Microsoft Photoshop and measured for plaque diameter. Small plaque morphology measured as <1 mm. Large plaque morphology measured as >1.5 mm. (A) WN-NY99, strain 382–99. (B) WNV 2002. (C) Bird 1153 (2003). (D) Bird 1171 (2003).

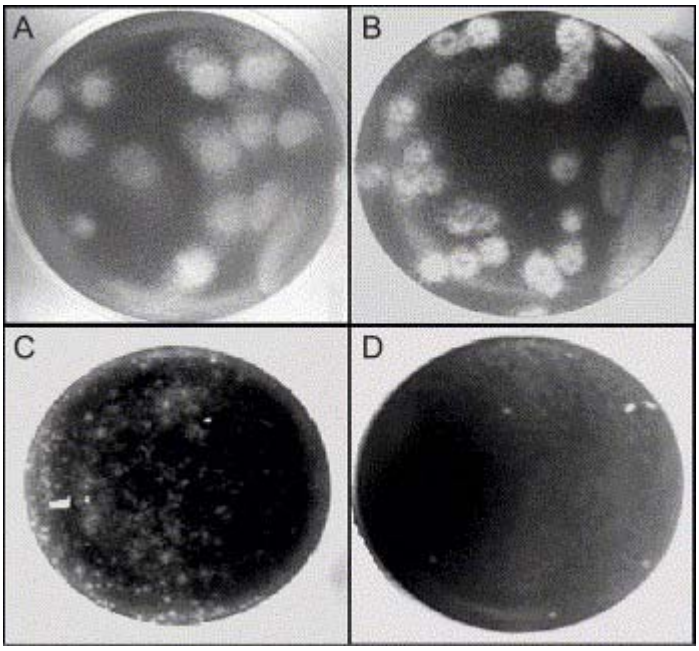
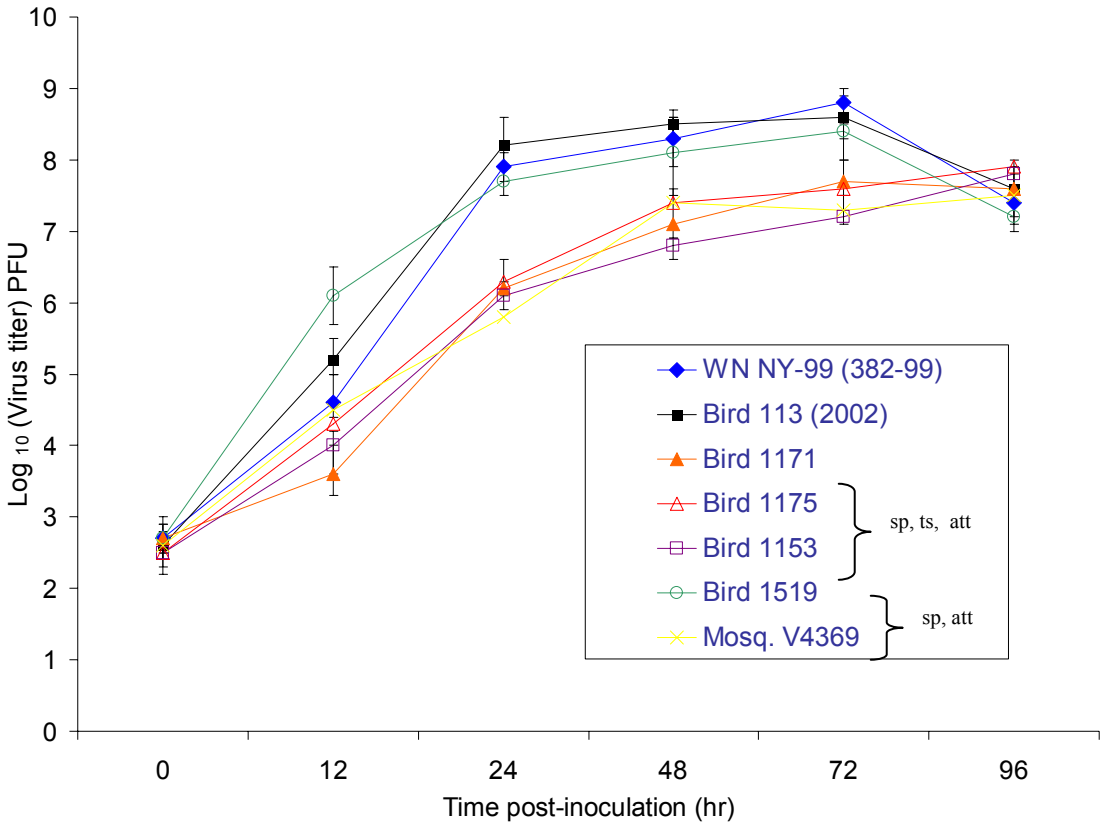


Fig. 5-2. Viral growth curve of 2003 sp WNV isolates in comparison to isolates from 1999 to 2002. Vero cells infected with WNV isolates at an MOI of 0.1 pfu/cell in triplicate in 12-well plates. Virus titers were then determined at indicated time points by plaque assay in Vero cells.



CHAPTER 6

MOLECULAR BASIS OF MOUSE ATTENUATION IN WEST NILE VIRUS VARIANTS

Introduction

The discovery of virus isolates with alterations in otherwise common phenotypic properties provides researchers with an opportunity to identify the molecular determinants that confer a given biological characteristic. Frequently, field isolation of viruses leads to the identification of genomic mutations that can be manipulated for vaccine development or for the design of improved therapeutics. For example, recent studies comparing the phenotypic properties of global field isolates of WNV have identified the importance of E protein glycosylation status as a virulence factor (Beasley et al., 2005). These studies are more straightforward when two viruses with similar genotypes possess marked differences in phenotype. The introduction of a WNV variant into North America with a highly mouse neuroinvasive phenotype has made it possible to elucidate subtle phenotypic variation in virus isolates that display a reduced capacity for neuroinvasion, at least in a mouse model. Because these isolates were made in the same

geographic region (even the same neighborhood) and during the same time period (within several days) as virus isolates possessing a highly mouse neuroinvasive phenotype, it was hypothesized that the mutations responsible for the alteration in phenotypic properties could be determined. Thus, this chapter describes full genome sequencing of variant isolates to identify potential mutations and the use of reverse genetics to confirm which mutations were responsible for the phenotypic changes in the Texas 2003 isolates.

Results

Nucleotide and deduced amino acid sequence comparisons

In order to determine those mutations in the WNV genome responsible for the phenotypic changes in West Nile viruses made in Texas during 2003, the complete viral genome of the 2003 WNV isolate Bird 1153 was sequenced (GenBank accession no. **AY712945**). Phylogenetic comparison of this sequence to all of the other complete genomes of WNV available from GenBank revealed that this isolate was a member of the North American clade of WNV isolates (Figure 4-2, Bird 1153 = Texas 2003 Harris 6 in the figure). Although there were a total of 36 nucleotide mutations in the genome of this virus relative to the prototype WN-NY99 strain, these mutations resulted in only 4 amino acid substitutions in the WNV polyprotein (prM-156, E-159, NS4B-249 and NS5-804) and 4 nucleotide substitutions in the 3'UTR (A10596G, C10774U, A10799G, and A10851G) (Table 6-1). One of the nucleotide changes in the 3'UTR (A10851G) was also identified in the WN-NY99 infectious clone (Beasley et al., 2005). In order to identify

nucleotide/amino acid substitutions shared by sp, ts, and/or attenuated isolates, each of the genes or untranslated regions of isolate Bird 1153 containing amino acid substitutions, or nucleotide mutations in the case of the 3'UTR, was sequenced in other isolates with a sp, ts, and/or attenuated phenotype. Table 6-1 shows the distribution of mutations among WNV isolates from 2002 and 2003 in comparison to WN-NY99 in those genes and untranslated regions where mutations were identified in the completely sequenced isolate. While several conserved mutations were identified among isolates from both 2002 and 2003, a single amino acid substitution (NS4B E249G) was found to be conserved in three isolates with a sp, ts, and mouse neuroinvasive attenuated phenotype (Bird 1153, Bird 1171, Bird 1175). Although this mutation was not identified in all attenuated 2003 isolates, it was hypothesized that the substitution in NS4B was responsible for the altered phenotype(s) of these three isolates. Comparison of this particular amino acid substitution with the analogous residue in other North American WNV isolates revealed that no other isolates described to date in GenBank contained this specific mutation.

Because the complete genome sequence of Bird 1153 did not reveal a single nucleotide/amino acid substitution shared only by sp and/or attenuated isolates, it suggested that multiple mutations were likely to be responsible for the range of phenotypes identified. Therefore, the complete genomes of three other 2003 WNV isolates (Bird 1461, GenBank accession no. **AY712947**; Bird 1171, GenBank accession no. **AY712946**; and Mosquito V4369, GenBank accession no. **AY712948**) were sequenced and compared to WN-NY99 (382-99) to identify additional mutations that

were likely to be responsible for the phenotypes (Table 6-1). With the exception of E V159A, nucleotide mutations in the genome of Bird 1461 (large plaque, non-ts, non-attenuated) encoded four unique amino acid substitutions compared with sp, ts, and/or attenuated isolates (NS3-E180D; NS3-E327K; NS4A-V134M; NS5-A618S). The genome sequence of Bird 1171 (sp, ts, and attenuated) was nearly homologous to that of Bird 1153 with two additional amino acid substitutions (NS5-R199L; NS5-A687D) and one additional substitution in the 3'UTR (G to U at residue 11000). Mosquito V4369 (sp, non-ts, attenuated) also shared the E V159A amino acid substitution and the 3'UTR nucleotide change at residue 10851, but this isolate revealed three additional amino acid substitutions (prM-N4D; NS4B-T240A; NS5-H295Y) and an additional nucleotide change in the 3'UTR (A to U at residue 10984). The prM-N4D substitution found in Mosquito V4369 was shared by another sp, non-ts, attenuated isolate, Bird 1519, suggesting its possible role in the altered phenotype(s) of these two isolates.

Site-directed mutagenesis of West Nile virus, WN-NY99, infectious clone and infectious virus rescue

A WNV infectious clone was obtained from Dr. Richard Kinney of CDC, Ft. Collins. The QuikChange® XL Site-Directed Mutagenesis Kit was used to mutate the pWN-CG plasmid containing the 3' half of the WNV genome to alter amino acid residue 249 of NS4B from Glu to Gly as described in Material and Methods (Figure 2-4). Infectious virus recovered from transfected Vero cells was aliquoted into ampoules

containing 0.5ml of supernatant. Viral RNA extracted from supernatant was amplified by RT-PCR using primers specific for the region containing the NS4B mutation to confirm the presence of the desired mutation. Plaque morphology and ts assays were performed and revealed the rescued virus to be a lp, non-ts variant (Fig. 6-1, Table 6-2). Lethality experiments comparing the NS4B E249G mutant to the NY99 infectious clone demonstrated that this mutant had the highly mouse neuroinvasive phenotype with an ip LD_{50} of 1.0 PFU.

Because the NS4B E249G mutant did not produce any of the phenotypes found in the field variants, additional mutant viruses were generated by incorporating several combinations of amino acid mutations and 3'UTR nucleotide mutations in both the pWN-AB and CG plasmids (i.e., 5' half and 3' half of the WNV genome, respectively (Table 6-2). In order to incorporate 3'UTR mutations found in the sp, ts, and attenuated isolates, a complete 3'UTR exchange was made between the NY99 infectious clone and Bird 1153, as described in Material and Methods. All of the engineered mutations and 3'UTR exchanges in the infectious clone pWN-AB and pWN-CG plasmids were confirmed by nucleotide sequencing of the modified plasmids in the regions containing the desired mutation(s). Transcription and transfection reactions were carried out as described in Materials and Methods. Sequences of primers used for mutagenesis are shown in Table 6-3.

Phenotypic characterization of WN-NY99 infectious clone-derived virus

Preliminary experiments revealed that the infectivity titers of virus derived from the infectious clone prior to site-directed mutagenesis were similar to those generated by inoculating cell cultures with wild-type WNV. Experiments also compared the mouse neuroinvasive/neurovirulence phenotypes of the parental strain WN-NY99 with virus derived from the infectious clone and found that both viruses have an ip LD_{50} of approximately 1.0 PFU and an ic LD_{50} ranging from 0.1 to 0.4 PFU (Table 6-2). Thus, the infectious clone derived virus is representative of the prototypical U.S. WNV strain from which the infectious clone was derived in its mouse neuroinvasive/ neurovirulence phenotype and plaque morphology (Figure 6-1) and serves as an appropriate genetic backbone with which to incorporate mutations of interest.

Phenotypic characterization of infectious clone derived mutants

A total of nine mutant viruses were made that incorporated mutations found in field isolates from Texas in 2003 that possessed phenotypic variation from the typical North American WNV (Table 6-2). Following rescue of infectious virus from transfected Vero cell cultures, viral supernatants were used for plaque titration, plaque morphology assays, and ts assays. All mutant viruses had plaque titers of $\geq 1.0 \times 10^4$ pfu/ml. Plaque morphology was visualized by crystal violet staining of 6-well plates after 3 dpi using

WN-NY99 to control for variation in plaque size. Two mutant viruses were found to have a sp phenotype (NS4B E249G + 3UTR and NS4B E249G + NS5A804V), while all others displayed a typical lp morphology (Fig. 6-1). Only mutants containing the 3'UTR from the sp, ts, att isolates (Bird 1153) were found to be ts, although the degree of temperature sensitivity at 41.0°C was not as significant in comparison to ts field isolates ($P < 0.05$ for each mutant ts isolate; $P < 0.001$ for ts field isolates versus non-ts plaque titer reduction by Student's t-Test).

Lethality experiments involving intraperitoneal inoculation in the Swiss Webster mouse model revealed that both sp mutant viruses (NS4B E249G + 3UTR and NS4B E249G + NS5A804V) were highly attenuated for mouse neuroinvasion (LD_{50} = >10,000 pfu and 2,000 pfu, respectively). Interestingly, incorporation of each point mutation alone, or even an entire exchange of the 3'UTR from the sp,ts, att isolate into the NY99ic, exhibited a highly neuroinvasive phenotype. All other mutant viruses produced ip LD_{50} survival time values similar to wild-type NY99 and the NY99 infectious-clone derived virus.

Multiplication kinetic studies

In order to measure the multiplication characteristics of the attenuated mutant viruses, Vero cells were infected in triplicate at an MOI of 0.1 pfu/cell in 12-well plates with NY99ic, NS4B E249G + 3'UTR (att), and NS4B E249G (non-att). Figure 6-2 shows the viral growth curve of each virus at various time points. While both NY99ic and NS4B

E249G derived viruses reached titers as high as 8.5 $\log_{10}$ pfu/ml, the sp, ts, att NS4BE249G + 3'UTR mutant failed to reach infectivity titers of greater than 7.0 $\log_{10}$ pfu/ml suggesting that this mutant has reduced levels of replication in Vero cells in comparison to prototypical WNV isolates. Additionally, there was an approximately 10-fold decrease in viral titers at many of the time points (12-72 hours), a finding that suggests inefficient replication at all time points post-infection.

Discussion

The emergence of phenotypic variants of WNV in Texas in 2003 led to an investigation to identify those mutations responsible and to understand how the mutations affected the phenotypic characteristics of WNV. By sequencing the complete genomes of several phenotypic variants of the North American WNV genotype and incorporating the identified mutations into the NY99 infectious clone, this study has identified several mutations in the WNV genome that reduce viral multiplication in vitro, confer temperature-sensitivity, and attenuate the ability of the virus to induce encephalitic disease in a mouse model. It was hypothesized that a single point mutation in the viral nonstructural protein NS4B would be responsible for imparting the sp, ts, and/or att WNV phenotype, however, this study has demonstrated that a combination of either single point mutations resulting in amino acid substitutions or nucleotide mutations in the 3'UTR were responsible for the previously described phenotypic variation. This study

demonstrates that a point mutation at residue 249 of the NS4B protein from a Glu to Gly in combination with a mutation in the NS5 protein at residue 804 (Ala to Val) or with three mutations to the viral 3'UTR (A10596G, C10774U, and A10799G) produce variants with sp, ts, and/or mouse attenuated phenotypes. Interestingly, the NS4B E249G mutation alone did not alter the phenotype of the infectious clone suggesting that the phenotypic variation observed in the field isolates was the result of more than one mutation to the viral genome. Additionally, evidence from ts mutants containing the 3'UTR exchange suggests that the ts phenotype is encoded by the 3'UTR but may be the result of more than a single nucleotide mutation in the 3'UTR. The requirement of multiple mutations to modify the phenotype of the North American WNV variants may explain why such variants are rarely isolated.

The NS4B protein of WNV is of unknown function and inferences concerning the influence of the E249G mutation remain speculative at best. However, previous studies of WNV and other closely related flaviviruses have shown this protein to be important to viral replication and pathogenesis. Westaway and others have described the accumulation of Kunjin virus NS4B protein the perinuclear region of infected cells and the ability of NS4B to translocate to the nucleus (Westaway et al., 1997). A live, attenuated vaccine strain (SA14-14-2) of Japanese encephalitis virus possesses an amino acid substitution in NS4B thought to be important in viral pathogenesis (Ni et al., 1995). Studies by Hanley and others have identified NS4B mutations that influence the susceptibility of SCID mice to dengue-4 virus infection (Hanley et al., 2003). Also, a mutation to the NS4B protein of dengue-2 virus has been associated with changes in the ability of the virus to inhibit

the interferon-signaling cascade, presumably by blocking STAT-1 phosphorylation (Munoz-Jordan et al., 2005). It is interesting to note that several Old World WNV isolates (e.g., Volgograd, Romania, Italy) contain different amino acid substitutions other than E to G at residue 249 of NS4B (Lanciotti et al., 2002). Also, a recent study by Rossi et al. (2004) identified the same mutation in a West Nile virus replicon that had established a persistent infection in different mammalian cell lines. Predictive structural models of this protein suggest that residue 249 is located in the carboxy-terminal portion of the protein in a region of the lumenal tail suggesting that this residue may form an interaction with components of the WNV replication complex or with other viral or cellular proteins (Fig. 6-3). Despite what little is known about the functional role of NS4B, amino acid substitutions to the protein may destabilize the conformation of the protein by disrupting bonds between amino acids important to protein stability or function. Future studies of the structure-function of the WNV NS4B protein will certainly help to define how mutations to this protein influence viral replication and pathogenesis.

Results from this study also suggest that mutations in the 3'UTR play a role in the generation of virus variants with a temperature-sensitive phenotype. All three mutant viruses exhibiting a ts phenotype possessed a 3'UTR gene swap from Bird 1153 (sp,ts,att). Genetic analysis showed that the 3'UTR contained three nucleotide mutations in comparison to NY99 each of which were located in or near the 5' stemloop structure (dumbbell-1; DB1) of the 3'UTR (Figure 6-4). The mapping of these mutations to a 3'UTR DB1 suggests their involvement in maintaining the secondary structure of the DB1, which in previous studies has been shown to be important for maintaining the

function of the 3'UTR (Tilgner et al., 2005). Studies have shown that deletions and mutations of conserved nucleotides in any one of the three stemloop structures predicted in the WNV 3'UTR reduced the replication efficiency of the mutant viruses produced (Brinton, 2002). In studies of dengue-4 and tick-borne encephalitis viruses, 3'UTR mutants have been highly attenuated with impaired growth characteristics (Proutski et al., 1999; Mandl et al., 1998). It has been proposed that mutations/deletions to the dumbbell structures lead to a structural rearrangement of the dumbbell directly or to pseudoknots located at the tips of the predicted dumbbells (See Figure 6-4). Consequently, it is hypothesized that this may disrupt the ability of the stemloop to stabilize and compartmentalize the replication complex during viral replication or to form binding sites for viral or cellular proteins important for replication or RNA synthesis (Brinton, 2002). Interestingly, the mutation at nucleotide 10799 is located in a region of the 5' dumbbell that is highly conserved in the *Flavivirus* genus, known as conserved sequence 2 (CS2). Although the function of this region remains unknown, it has been suggested that the CS2 region is involved in viral replication (Markoff, 2003). Mutation of a C to U at nucleotide 10774 is also of interest because of the possibility that this change may influence proper base pairing and folding patterns of the DB1 stemloop structure. The mutation at nucleotide 10596 is less likely to influence the structure of the 3'UTR because it is found outside of the conserved stemploop region. Regardless, the ts phenotype of isolates containing these mutations may be the result of alterations to the 3'UTR secondary structure rendering the dumbbell structure sensitive to high temperature conditions. Under these conditions, it is possible that the function of the dumbbell region is impaired,

resulting in reduced replication. While the mutations in the 3'UTR may influence temperature-sensitivity, the 3'UTR mutations alone did not attenuate the infectious clone-derived virus, suggesting that these mutations alone do not alter viral replication under the physiological conditions of the mouse model. At present this is a hypothesis as the growth kinetics of the 3'UTR mutant alone have not been investigated. Additionally, the NS4B E249G mutation did not alter the viral phenotype, suggesting that these mutations act in concert to significantly alter the biological properties of the virus.

In addition to mutations in the 3'UTR, a sp and mouse attenuated mutant was generated with a substitution at residue 804 in the NS5 protein along with the NS4B substitution. The NS5 gene of WNV encodes a single protein believed to possess both an N-terminal region with methyltransferase activity and a C-terminal region with RNA-dependent RNA polymerase (RdRp) activity. The NS5 A804V mutation is located in the C-terminal region of the protein, but is found outside of any of the conserved motifs previously described as important to RdRp activity (Ackermann and Padmanabhan, 2001). The presence of a mutation in this region that results in a sp and mouse attenuated viruses suggests the possibility that the Ala at residue 804 is important for the activity of the RdRp. This hypothesis warrants further investigation as the growth kinetics of the NS5 A804V mutant in cell culture have not been performed. More detailed analyses of the consequences of this mutation on polymerase function are required.

Despite the identification of several multi-site mutations that impart the sp, ts, and mouse attenuated phenotypes observed in WNV field isolates, it is clear from genomic sequencing of variant viruses (e.g., Mosq. v4369) that additional mutations resulting in sp

and attenuated mutants were not found. Even though several mutations from the Mosq. v4369 were incorporated into the NY99ic, no mutant viruses were produced that had a sp or attenuated phenotype. Additional studies of infectious clone mutants will be required to further characterize other mutations in the viral genome that may result in phenotypic variation in WNV.

Table 6-1. Nucleotide changes and deduced amino acid sustitutions of 2002-2003 Texas WNV isolates compared to WN-NY99 (AF196835).

WNV isolate	Nucleotide position	Gene/region	Nucleotide change	Amino acid change
Harris Co., TX 2002	1442	E	U to C	V159A
lp, non-ts, non-att	7699	NS5	A to C	T6P
	10408	3' UTR	C to U	n/a
	10851	3' UTR	A to G	n/a
Bolivar P., TX 2002*	1192	E	A to G	T76A
lp, non-ts, non-att	2749	NS1	A to G	E94G
	3937	NS2A	G to A	V138I
	7432	NS4B	G to A	V173I
	9256	NS5	C to U	T526I
	10494	3' UTR	U to C	n/a
	10768	3' UTR	U to A	n/a
	10851	3' UTR	A to G	n/a
Bird 1461*	1442	E	U to C	V159A
lp, non-ts, non-att	5151	NS3	A to U	E180D
	5593	NS3	G to A	E327K
	6871	NS4A	G to A	V134M
	9535	NS5	G to U	A618S
	10408	3' UTR	C to T	n/a
	10851	3' UTR	A to G	n/a
Bird 1153*	931	prM	G to A	V156I
sp, ts, att	1442	E	U to C	V159A
	7661	NS4B	A to G	E249G
	10091	NS5	C to U	A804V
	10596	3' UTR	A to G	n/a
	10774	3' UTR	C to U	n/a
	10799	3' UTR	A to G	n/a
	10851	3' UTR	A to G	n/a
Bird 1171*	931	prM	G to A	V156I
sp, ts, att	1442	E	U to C	V159A
	7661	NS4B	A to G	E249G
	8279	NS5	G to U	R199L
	9743	NS5	C to A	A687D
	10091	NS5	C to U	A804V
	10596	3' UTR	A to G	n/a
	10774	3' UTR	C to U	n/a
	10799	3' UTR	A to G	n/a
	10851	3' UTR	A to G	n/a
	11000	3' UTR	G to U	n/a
Bird 1175	1442	E	U to C	V159A
sp, ts, att	7661	NS4B	A to G	E249G
	10408	3'UTR	C to U	n/a
	10851	3' UTR	A to G	n/a
Bird 1519	478	prM	A to G	N4D
sp, att	1442	E	U to C	V159A
	10851	3' UTR	A to G	n/a
Mosq. V4369*	478	prM	A to G	N4D
sp, att	1442	E	U to C	V159A
	7636	NS4B	A to G	T240A
	8566	NS5	C to U	H295Y
*Indicates isolate has been completely sequenced	10851	3' UTR	A to G	n/a
	10984	3' UTR	A to U	n/a

Table 6-2. Mutant infectious clone derived viruses and their phenotypic properties.

Virus	Plaque size	Temperature sensitivity			Mouse neuroinvasiveness and neurovirulence			
					Intraperitoneal inoculation		Intracerebral inoculation	
		37°C[a]	41.0°C[a]	Δ	LD50 (PFU)	AST ± s.d.	LD50 (PFU)	AST ± s.d.
WN-NY99 (382-99)	lp	8.1	7.6	0.5	0.8	8.0 ± 1.2	0.1	6.4 ± 0.9
WN-NY99 ic	lp	7.0	7.3	0.3	1.0	7.2 ± 0.6	0.4	6.0 ± 0.2
Bird 1153	**sp**	7.7	4.8	***2.9****	**23,000**	n/a	0.1	6.7 ± 0.3
Mosq. V4369	**sp**	7.2	6.8	0.4	**645,000**	n/a	0.2	6.2 ± 0.3
NS4B E249G (sp, ts, att)	lp	6.2	6.0	0.2	1.2	7.5 ± 0.4	0.1	6.0 ± 0.2
prM V156I (sp, ts, att)	lp	5.9	6.1	0.2	0.7	8.6 ± 1.4	n.d.	n.d.
3'UTR (sp, ts, att)	lp	6.4	5.0	***1.4****	0.6	8.8 ± 2.0	n.d.	n.d.
prM V156I + NS4B E249G (sp, ts, att)	lp	7.0	7.2	0.2	1.4	8.8 ± 1.5	n.d.	n.d.
prM V156I + 3'UTR (sp, ts, att)	lp	6.6	5.1	***1.5****	4.2	9.0 ± 1.0	n.d.	n.d.
NS4B E249G + 3'UTR (sp, ts, att)	**sp**	5.0	3.2	***1.8****	**>10,000**	n/a	n.d.	n.d.
NS4B E249G + NS5 A804V (sp, ts, att)	**sp**	6.2	6.4	0.2	**2,000**	9.2 ± 1.2	n.d.	n.d.
prM N4D (sp, att)	lp	7.3	7.1	0.2	1.2	9.4 ± 1.6	n.d.	n.d.
NS4B T240A + NS5 H295Y (sp, att)	lp	5.9	6.1	0.2	0.7	7.3 ± 0.5	n.d.	n.d.

Small plaque (sp) morphology = < 1.0 mm (shown as bold/underlined)

Large plaque (lp) morphology = > 1.5 mm

[a] Log_{10} Plaque titer at each temperature (pfu/ml)

Δ = difference in plaque titer at 37.0 and 41.0

ts = temperature-sensitivite phenotype at 41.0°C (shown as bold/italics)

att = attenuated for mouse neuroinvasiveness (shown in bold)

3'UTR = A10596G, C10774U, A10799G, and A10851G

P values were calculated versus average non-ts plaque titer reduction by Student t-Test; *Indicates statistical significance.

Table 6-3. Mutagenic Primers used for site-directed mutagenesis

NS4B E249G (sp, ts, att)
7645 For
CTCATAAAGAACATGGGAAAACCAGGACTAAAAAGAGGTGGGGC
7689 Rev
GCCCCACCTCTTTTTAGTCCTGGTTTTCCCATGTTCTTTATGAG

prM V156I (sp, ts, att)
915 For
GCAGAGAGTTGTGTTTATCGTGCTATTGCTTTTGGTGGCCCCAGC
959 Rev
GCTGGGGCCACCAAAAGCAATAGCACGATAAACACAACTCTCTGC

prM N4D (sp, att)
451For
GCCAGCGTAGGAGCAGTTACCCTCTCTGACTTCCAAGGGAAGG
493 REV
CCTTCCCTTGGAAGTCAGAGAGGGTAACTGCTCCTACGCTGGC

NS4B T240A (sp, att)
7636 For
GGT TGG TTG TCA TGT CTA TCC ATA GCA TGG ACA CTC
7636 Rev
GAG TGT CCA TGC TAT GGA TAG ACA TGA CAA CCA ACC

NS5 H295Y (sp, att)
8566 For
CGT GAG TAC AGT TCG ACG TGG CAC TAC GAT GAG AAC CAC CC
8566 Rev
GG GTG GTT CTC ATC GTA GTG CCA CGT CGA ACT GTA CTC ACG

NS5 A804V (sp, ts, att)
10091 For
CCA CGT GGT CCA TCC ATG TAG GAG GAG AGT GG
10091 Rev
CC ACT CTC CTC CTA CAT GGA TGG ACC ACG TGG

Figure 6-1. Plaque morphology of NY99 infectious clone mutants

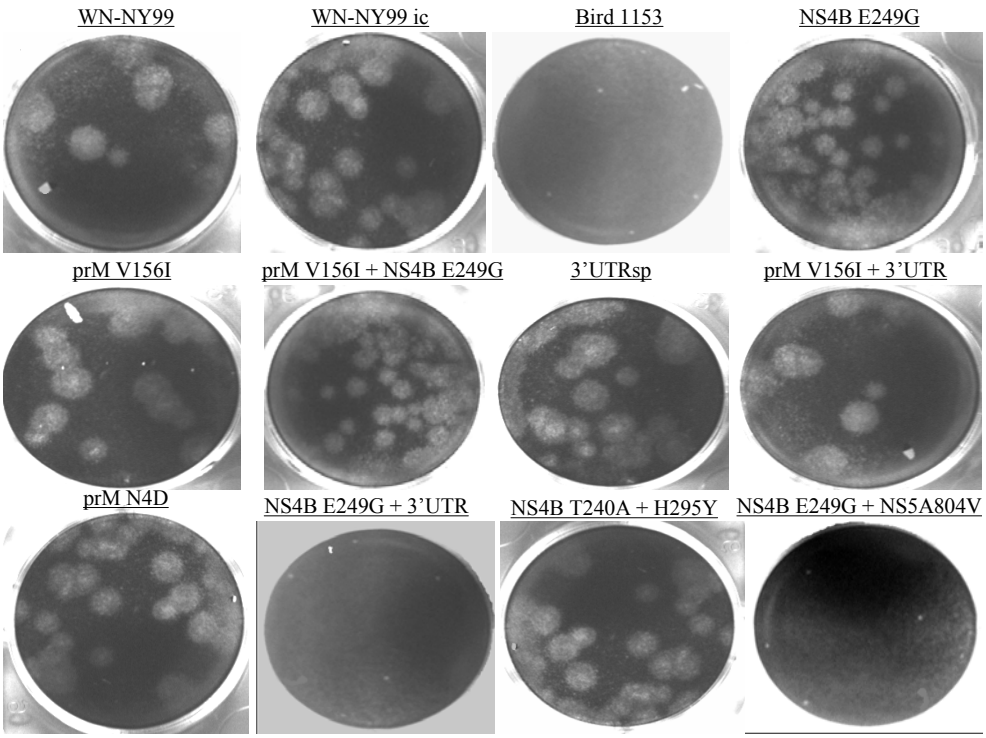


Figure 6-2. Viral growth curve of attenuated infectious clone mutants in comparison to NY99ic and a non-attenuated infectious clone mutant.

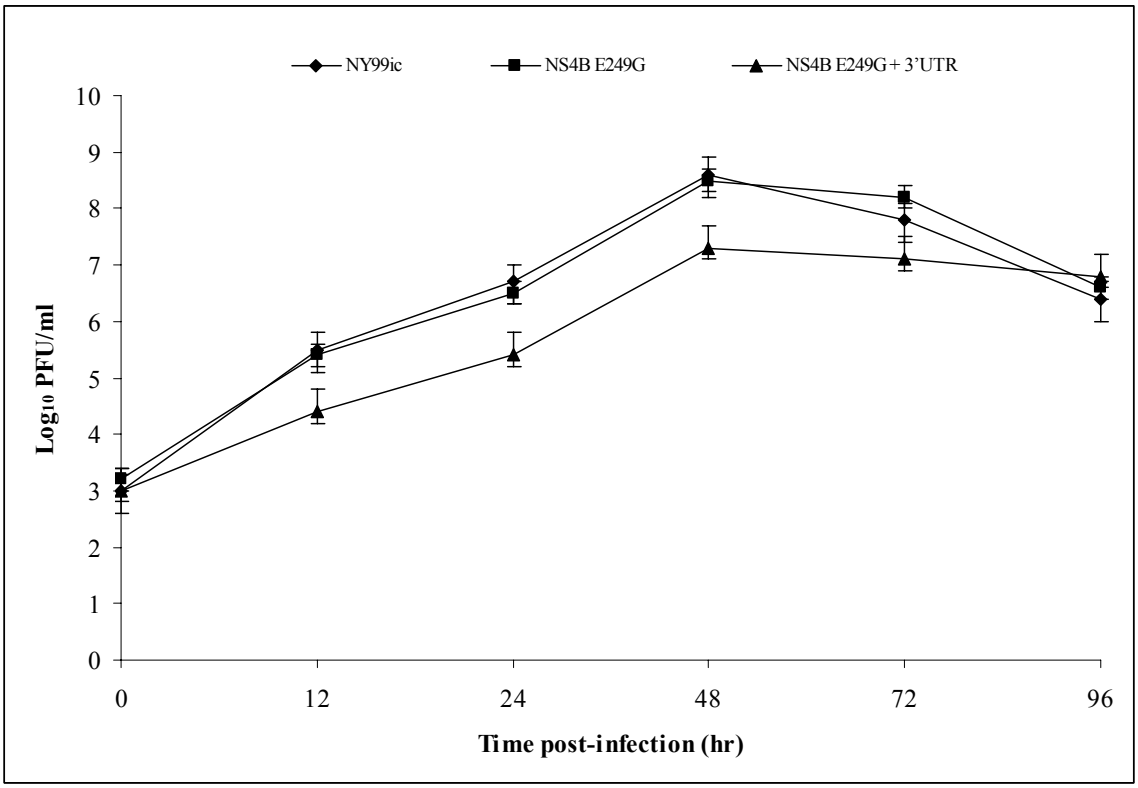


Figure 6-3. Predicted hydrophobicity plot of the NS4B protein generated by the SOSUI program. NS4B E249G (denoted by arrow) was predicted to reside in the lumenal C-terminal tail of the WNV NS4B protein (Courtesy of J. Wicker, unpublished).

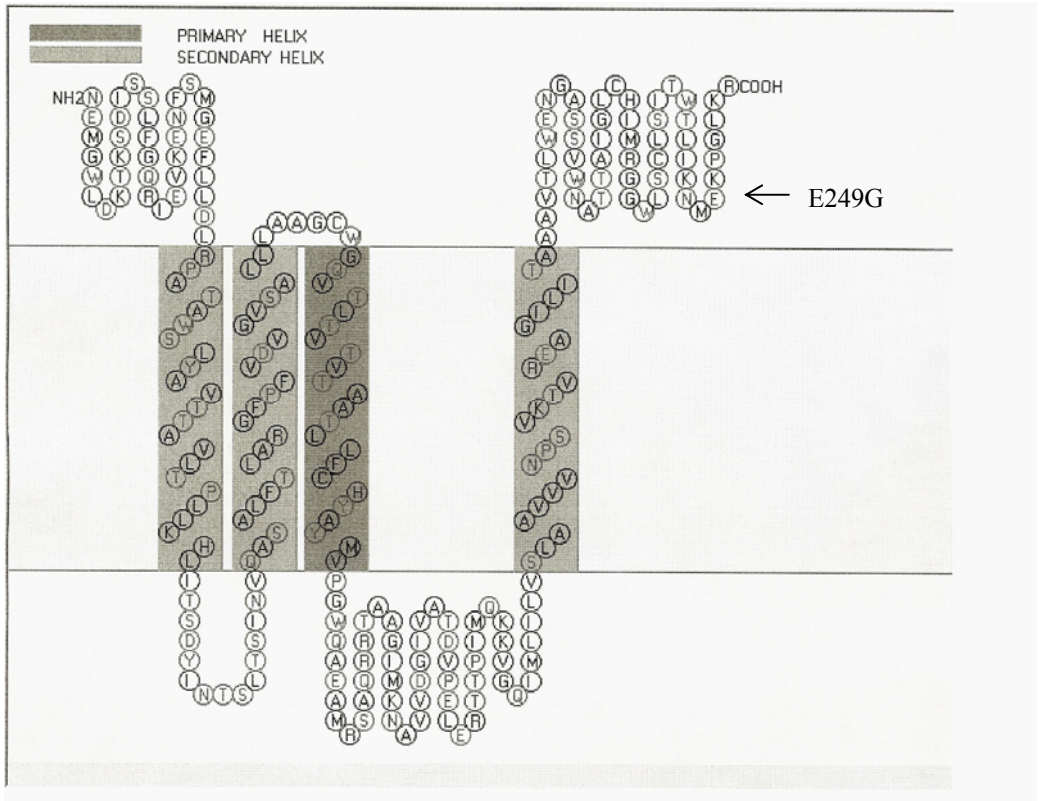
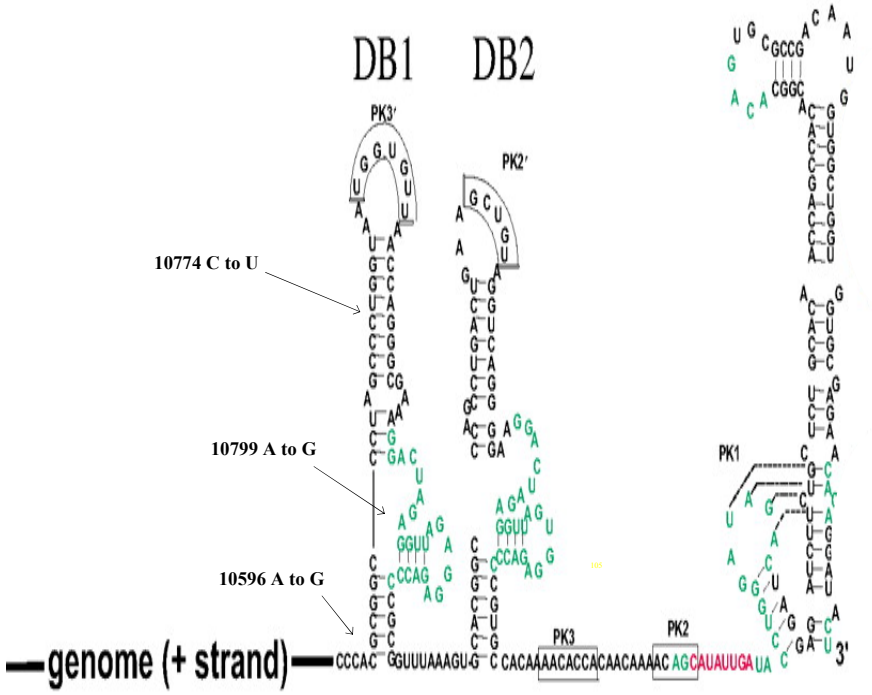


Figure 6-4. Predicted secondary structure of WNV 3'UTR showing location of nucleotide mutations in sp, ts, att isolates. DB1 and DB2: dumbbell 1 and 2 (Adapted from Brinton, 2002).



CHAPTER 7

DISCUSSION

The advent of nucleic acid sequencing and advances in sequencing technology have greatly contributed to the ever growing field of molecular epidemiology, providing tools that allow the comparison of entire viral genomes rapidly and precisely. In addition, genetic engineering, and in particular reverse genetics, has given researchers a tool to accurately access the biological significance of genetic polymorphisms in experimentally controlled systems. The fusion of these two fields has expanded the ability of researchers to address questions pertaining to the relationship between public health and basic mechanistic biology. This dissertation has taken advantage of advances in both fields to examine how the introduction of a virus to a new environment affects the genetic composition of the viral population at large, the biological properties of phenotypic variants, and the polymorphisms that give rise to those variants.

In the first aim of this dissertation, studies to determine the extent of the genetic divergence of WNV since its introduction into North America identified several genetic variants in the North American population providing evidence of the importance of both resident and migratory birds in the dispersal of WNV and the tight geographical and temporal associations among distinct genotypes. Moreover, these studies led to the

discovery of a dominant genotype, whose emergence appears to have led to the systematic displacement of earlier virus genotypes once found throughout the eastern U.S. The second aim of this research proceeded to examine the consequences of the emergence of genetic variants by considering potential changes in viral phenotype. As a result, these studies demonstrated the existence of phenotypically distinct viruses that expressed very different biological properties from the prototypical West Nile viruses that were circulating in North America. The discovery of WNV variants with mouse attenuated phenotypes was the first evidence of phenotypic variation described in the North American WNV population supporting the suggestion and possibility that the virus may eventually become adapted its new environment by, possibly, evolving towards a less virulent phenotype. Finally, utilizing a combination of genomic sequencing and reverse genetics technology, experiments were conducted using a WNV infectious clone to precisely identify mutations in the North American WNV genome that gave rise to the phenotypic variants. Accordingly, the mapping of amino acid substitutions to viral proteins and nucleotide mutations to the viral 3' UTR were shown to significantly alter several of the biological properties of the virus. These investigations illustrated novel molecular determinants of WNV attenuation in the nonstructural proteins and 3' UTR and the propensity for multiple mutations to act in combination to modify the phenotypic characteristics of WNV. This chapter discusses the implications of the findings of this dissertation and compares these findings to other studies that pertain to the molecular epidemiology of flaviviruses.

Although WNV has only been endemic in the U.S. for six years, continued multiplication of the virus population has led to the accumulation of mutations in many portions of the genome and the emergence of a dominant genotype throughout North America. Like other RNA viruses, WNV has a relatively high mutation rate during genome replication due to its utilization of the viral RNA-dependent RNA polymerase (RdRp). The RdRp of WNV, like other flaviviruses, is responsible for the accumulation of large numbers of misinsertion errors during replication (in the range10^{-3} to 10^{-5} substitutions per nucleotide and per round of copying). Also, the polymerase is not believed to contain a domain with 3' to 5' exonucleolytic proofreading activity. In addition, mismatch repair mechanisms are unlikely to operate on replicating RNA and cannot operate on single-stranded RNA progeny genomes (Holland, 1996). The error-prone nature of the flavivirus RdRp results in rapid genetic divergence because of the inability of the polymerase to proofread for and correct nucleotide mutations made in the viral genome during replication. While some members of the family, such as hepatitis C virus (HCV), accumulate a large percentage of mutations in the viral genome during infection in a single host, others accumulate much fewer (Duarte et al., 1994). Studies of mosquito-borne viruses, in particular, have shown that they accumulate nucleotide mutations at a slower pace than non-mosquito-borne viruses that infect only one host. Experiments conducted using vesicular stomatitis virus (VSV) (Novella et al., 1995) and eastern equine encephalitis (EEE) virus (Weaver et al., 1999) have addressed the question of the evolutionary constraints placed on mosquito-borne viruses and suggest that because mosquito-borne viruses must replicate in two very different selective

environments (i.e., invertebrate and vertebrate hosts), the viral genomes that accumulate large numbers of mutations become less fit (debilitated) or non-replicative allowing only those genomes with few mutations to be transmissible between altering host environments. Thus, even though WNV has accumulated mutations to the genome, it is possible that the conserved accumulation of only a few nucleotide mutations and a single amino acid substitution (Table 4-4) is a consequence of constraints placed on the WNV genomes that allow them to remain fit and replication competent.

Selective pressures, such as the continued capacity to replicate in both arthropod and vertebrate hosts and/or the escape from host immunity (immune selection), forcing viral evolution are often cited as primary factors in the evolution of flaviviruses. Studies of DEN viruses by Grenfell et al. (2004) suggest that evolution of DEN virus at the level of the E protein and possibly other viral proteins is driven by immune selective pressures that force the divergence of DEN strains as a means of avoiding neutralization by the immune system of the host. Under this type of selection, these studies suggest that selective pressures lead to a bottleneck of viral genotypes resulting in the emergence of a predominant genotype. While immune selection may eventually drive the divergence of WNV as it continues to infect more and more birds, results from this study do not reveal the accumulation of a large number of amino acid substitutions in the E protein and none in domain III of the E protein (the putative receptor-binding domain and primary immunogenic domain) (Table 4-3). However, other selective pressures, such as the two-host constraint, may have contributed to the accumulation of specific mutations in the viral genomes studied in Chapters 3 and 4. Evidence that selective pressures may

influence the divergence of WNV is related to the accumulation of mutations in some genes but not others, as discussed in Chapter 4. The ability of some genes to withstand mutations, while others cannot, may reflect a specific pressure imposed on the function of a given protein. As more sequence data become available, it may be possible to infer if selection pressures are acting on certain genes or untranslated regions by analyzing the ratio of nonsynonymous to synonymous substitutions. Additionally, as the WNV sequence database expands over time, it will be possible to calculate the rates of evolutionary change incurred by the North American WNV population.

Although selective pressures may have contributed, the genetic divergence of WNV isolates studied in this dissertation may simply be the consequence of genetic drift in which mutations in the genome result in a random pattern due to mutations during viral replication. While genetic drift may have resulted in the early genetic variation observed in WNV isolates, results from Chapter 4 demonstrating the emergence of a dominant North American variant, suggest that at least some form of selective pressure has been responsible for the accumulation of 13 nucleotides in the dominant genotype. The identification of this genotype in all isolates made west of the Mississippi River after 2002, suggest that the emergence of this genotype was driven by a selective advantage, either genetic or phenotypic, resulting in the subsequent spread of only the dominant genotype across North America. Regardless of the selective pressure acting on North American WNV, the correlation between distinct genetic variants and geographical location of isolates has proven useful in better understanding the role of birds in the dispersal of WNV. Results from Chapter 3 and 4 illustrate the importance of both

resident and migratory birds in the dispersion of virus in the U.S., Canada, and Mexico. The current lack of isolates, and consequently sequence information from regions outside of North America, especially in the Caribbean islands and southward into Central America, has made it difficult to determine the influence that migratory birds have in the long distance dispersal of WNV into new regions. Clearly, additional isolates from Central and South America will address this question in the future. Also, it will be important for future studies of WNV molecular epidemiology to continue to acquire sequence information from regions in which sequence data already exist in order to continue to measure genetic divergence of viruses in endemic foci. In doing so, it may be possible to determine if the rate of genetic divergence (and the selective pressures acting on a virus population) differs between endemic regions and new areas of virus introduction. Finally, results from this dissertation stress the importance of virus isolation from different locations and different years to provide a source of virus variants (genetic or phenotypic) that can be used to study the spatial and temporal evolution of a virus.

For many years, molecular epidemiological studies of flaviviruses have attempted to answer questions about how genetic variation leads to changes in the ecology and epidemiology of a given virus. Studies of the molecular epidemiology of the four DEN viruses have provided a large amount of data from which to draw conclusions concerning how viral evolution affects epidemiology. Sequencing of DEN-2 viruses isolated directly from patients in both Southeast Asia and the Americas suggest a geographical correlation between high virulence strains (Asian) and low virulence strains (American) (Leitmeyer et al., 1999). This study also showed that the virulence differences were consistently

associated with alterations in the untranslated regions of the genomes and with charge differences between amino acids in the prM, E, NS4B, and NS5 genes. Other studies have shown that 5'UTR and 3'UTR exchanges of the American untranslated regions into an infectious clone of an Asian strain greatly reduce the levels of replication in human primary dendritic cells (Cologna and Rico-Hesse, 2003), suggesting virulence differences between these strains. Additionally, studies demonstrating variation in the transmissibility of different DEN virus strains by mosquitoes provide evidence that some strains may be of greater epidemiological significance (Bosio et al., 2000).

It is also clear from phylogenetic analyses of DEN 2 and DEN 3 viruses in the Americas (Rico-Hesse et al., 1997; Watts et al., 1999; Usuku et l., 2001) that the displacement of one genotype by another has occurred in many countries in Central and South America, often resulting in increased cases of dengue hemorrhagic fever (DHF). Finally, the introduction or emergence of a DEN genotype into a region with a different genotype has been shown to lead to the extinction (local or complete) of the original genotype (Holmes and Twiddy, 2003; Wittke et al., 2002). Whether or not this relates to fitness advantages of one genotype versus another has yet to be clearly addressed. Findings from this dissertation have revealed the displacement of one genotype (Eastern U.S. genotype) by another genotype (North American 2002-2004 dominant genotype) and suggest that earlier genotypes are now extinct (Eastern U.S. and southeast Texas genotypes). Ebel et al. (2004) suggest that this displacement/extinction may be related to increased transmission efficiency of the dominant genotype by *Culex* mosquitoes. Additional mosquito infectivity studies comparing viruses from different genotypes may

help to assess this in greater detail. Future studies should also attempt to determine if other fitness advantages exist between the dominant genotype and other North American WNV genotypes. It should be possible to address this question experimentally using competition assays to compare in vitro fitness between isolates belonging to different genotypes.

Because many molecular epidemiological studies rely merely on the correlation between genetic variation and a given epidemiological variable, it is often difficult to test the correlation using experimental methods. Recent technological breakthroughs in recombinant engineering, however, now allow researchers to test these correlations in more well defined experimental systems. The third aim of this dissertation was possible because of the ability of a WNV reverse genetics system to experimentally test the association between certain viral genotypes and the phenotypic properties that they displayed.

The results from Chapter 6 have identified mutations in the WNV genome that may help to explain the function of several of the WNV proteins in terms of virus virulence, such as NS4B or NS5, as well as the structural and functional implications of mutations in the viral 3'UTR. With the identification of several molecular determinants of mouse attenuation, future studies will no doubt shed light on the molecular mechanisms responsible. Although the function of NS4B is currently unknown, it remains to be seen how the NS4B E249G mutation directly affects the replication of the virus in terms of RNA synthesis, transcription, genome replication, or a number of other potentially important functions. Single-step growth curves in cell culture should be

undertaken to measure the replication kinetics of viruses with this, and other mutations. Mutant viruses should also be characterized in vitro with regard to their RNA and protein content to examine potential deleterious affects of mutations on transcription and translation. Additionally, experiments to determine the affect of the NS4B E249G mutation on the ability of viruses to inhibit the interferon-signaling cascade, as has been found for some NS4B DEN 2 mutant viruses (Munoz-Jordan et al., 2005), may also reveal an important functional component of the NS4B E249 residue. Further structural modeling of the viral NS5 protein may help to explain how the NS5 V804A substitution affects protein function, and if this affect is a consequence of viral polymerase dysfunction. In vitro polymerase activity assays may also reveal an important role of this mutation to NS5 function. The identification of three mutations in the 3'UTR of WNV variants that contribute to the sp, ts and attenuating phenotype also calls for future studies to identify exactly which mutations affect viral phenotype or if all are necessary. Based on the position of each of the 3'UTR mutations relative to conserved secondary structures, it is likely that the mutations found directly within dumbbell 1 of the 3'UTR (C10774U and A10799G) are responsible for the observed phenotype (Fig. 6-4). Site-directed mutagenesis of each of the individual 3'UTR mutations alone and in various combinations should identify more precisely the mutations in the 3'UTR that influence viral phenotype. This may also help to explain how secondary structures in the 3'UTR influence viral replication and other components of the viral life cycle.

Instrumental to the success of aim 3 was the use of a WNV infectious clone based on the genome of WN-NY99. Not only are these types of reverse genetics studies useful

in addressing molecular mechanistic questions, they also have the potential to be exploited for improved vaccine design. Today, many researchers rely on reverse genetics technology in order to develop novel vaccines (Hoffman et al., 2002; Lai and Monath, 2003; Nalca et al., 2003). As discussed in the introduction, at the forefront of human WNV vaccine development is a vaccine that uses infectious clone technology to incorporate the prM and E protein genes into a YF virus 17D backbone (Lai and Monath, 2003). Similarly, site-directed mutagenesis of infectious clone plasmids has the potential to introduce attenuating mutations into the genome of a virus for vaccine development. The results from Chapter 6 have identified three potential protein genes/untranslated region (NS4B, NS5, and 3'UTR) with which to incorporate attenuating mutations in order to enhance the attenuating features of a recombinant vaccine candidate or to add additional mutations to an existing attenuated genome to facilitate the stability of the attenuation phenotype and to help avoid the possibility of reversions of the vaccine. It will be important to determine if the mutations identified by this dissertation have similar affects in other flaviviruses that could potentially be useful for vaccine design. Also, if indeed these mutations were to be incorporated into a vaccine candidate, future studies would be necessary to determine if the attenuation of neuroinvasiveness identified in the mouse model correlates with attenuation in other animal models of WNV, such as birds, horses, or non-human primates. Additional studies concerning the adaptive and innate immune responses following infection by a vaccine candidate containing these mutations would also be critical to evaluating their usefulness for vaccine design.

It is unlikely that we will ever know how WNV was introduced into North America, although it was most likely related to a human-mediated event, such as a stow-away bird or mosquito on an airplane or boat or an infected human or horse. If migratory or storm displaced birds were the culprit, the introduction of WNV into North America would have likely occurred more than once and more recently in history. Irrespective of this, the introduction of WNV into North America underscores the propensity for human activities to introduce disease agents into new environments. It may only be a matter of time before other mosquito-borne viruses, such as JE virus, find their way into North America. Another question which can only be answered in due time relates to the future epidemic potential of WNV in the Western Hemisphere. While the human disease incidence appears to be decreasing over time in the U.S. and Canada, and appears to be practically non-existent in Mexico, it remains to be seen whether or not WNV will cause significant disease outbreaks in Central and South America, especially as the virus begins to circulate in regions that have no endemic flaviviruses, and hence, no potential cross-protective immunity. Additionally, it is unclear at this point in time if WNV will continue to cause sporadic outbreaks of WN fever and encephalitic disease in humans or if the virus will become more enzootic over time.

Other close relatives of WNV, namely SLE and JE viruses, have extremely different disease potential in the regions in which they are endemic. SLE is distributed from southern Canada to Argentina and is maintained in a transmission cycle similar to WNV (although it is not normally associated with avian mortality). Between 1930 and 1975 there were major outbreaks of disease in the U.S. that mostly occurred in regions

where conditions allowed for large mosquito populations and increased human contact. During this time, there were over 1,000 deaths and many thousands of cases of severe illness attributed to SLE virus infection. Since 1975, however, there have been only 1,400 cases reported to the CDC (approximately 57.4 cases per year) with the majority of these cases occurring in specific cities or regions within a state (Reisen, 2003). Presumably this change in disease incidence relates to increased use of air conditioning and behavioral changes in humans that have reduced mosquito contact. JE virus is dispersed throughout Asia and has most recently been detected in northern Australia. Like SLE and WN viruses, JE is transmitted from birds to mosquitoes in an enzootic transmission cycle, but has also been found to be amplified in nature by pigs as well. Unlike SLE virus, however, JE virus is a major cause of viral encephalitis every year in many Asian countries resulting in an estimated 30-50,000 cases per year with approximately 30% of those cases involving neurological illness. The case-fatality rate is also extremely high and has been reported to be >20% in many countries (Endy and Nisalak, 2002).

Since the introduction of WNV into North America, the disease incidence has been more comparable to the incidence of JE virus in Asia than to SLE in North America, with over 16,500 cases of disease reported since its introduction and over 600 deaths (CDC, Arbonet). While this high incidence of disease may be the result of introduction of the virus into an immunologically naïve population, it remains to be seen how the pattern of disease incidence will shift now that the virus has become endemic across the continent. As is already evident from case reports in 2004 and 2005, the number of cases and deaths has begun to decline, perhaps due to regional ecological changes, such as

decreased rainfall and mosquito abundance or a decline in reservoir bird populations. Alternatively, protective immunity in the bird population may be reducing the transmission of the virus in its natural cycle. Human intervention in the form of mosquito control measures and heightened person protection to avoid mosquito contact may also be factors in the decline in case numbers. It may also be possible, although not evident from this dissertation, that as the virus has become established, it has begun to adapt to its environment by becoming less virulent. Although results from this dissertation have identified mouse attenuated variants in 2003, the majority of WN viruses circulating in North America retain the highly mouse neuroinvasive phenotype of the progenitor WN-NY99 (this dissertation). Because the strain of WNV introduced into North America was a strain capable of causing severe illness and neurological disease, it will most likely continue to do so in the coming years. Whether or not we will continue to see disease outbreaks as intense as those occurring in 2002 and 2003 is unclear. And while WNV will no doubt remain endemic in North America, only time will tell if the virus becomes more enzootic in its transmission, like SLE virus, with only rare cases each year, or if WNV will continue to result in annual outbreaks over a wide geographical region that coincide with peaks in mosquito density. Regardless, the emergence and spread of WNV into North America will continue to provide arbovirologists a unique opportunity to learn more about the ecology and epidemiology of an emerging arboviral disease.

REFERENCES

Ackermann M and Padmanabhan R. 2001. De novo synthesis of RNA by the dengue virus RNA-dependent RNA polymerase exhibits temperature dependence at the initiation but not elongation phase. Journal of Biological Chemistry 276: 39926-39937.

Amberg SM and Rice CM. 1999. Mutagenesis of the NS2B-NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. Journal of Virology 73: 8083-8094.

Anderson JF, Andreadis TG, Vossbrinck CR, Tirrell S, Wakem EM, French RA. 1999. Isolation of West Nile virus from mosquitoes, crows, and a Cooper's hawk in Connecticut. Science 286: 2331–3.

Anderson JF, Vossbrinck CR, Andreadis TG, Beckwith WH, Mayo DR. 2001. A phylogenetic approach to following West Nile virus in Connecticut. Proceedings of the National Academy of Sciences 98: 12885–12889.

Arroyo J, Miller C, Catalan J, Myers GA, Ratterree MS, Trent DW, Monath TP. 2004. ChimeriVax-West Nile virus live-attenuated vaccine: preclinical evaluation of safety, immunogenicity, and efficacy. Journal of Virology 78: 12497-12507.

Bakonyi T, Hubalek Z, Rudolf I, Nowotny N. 2005. Novel flavivirus or new lineage of West Nile virus, central Europe. Emerging Infectious Disease 11: 225-231.

Beasley DWC, Davis CT, Estrada-Franco J, Navarro-Lopez R, Campomanes-Cortes A, Tesh RB, Weaver SC, Barrett ADT. 2004. Complete genome sequence of a West Nile virus isolate from Mexico and identification of an attenuating mutation. Emerging Infectious Disease 10: 2221-2224.

Beasley DWC, Davis CT, Estrada-Franco J, Navarro-Lopez R, Campomanes-Cortes A, Tesh RB, Weaver SC, Barrett ADT. 2004. Genome sequence and attenuating mutations in West Nile virus isolate from Mexico. Emerging Infectious Disease 10: 2221-24.

Beasley DW, Whiteman MC, Zhang S, Huang CY, Schneider BS, Smith DR, Gromowski GD, Higgs S, Kinney RM, Barrett AD. 2005. Envelope protein glycosylation status influences mouse neuroinvasion phenotype of genetic lineage 1 West Nile virus strains. Journal of Virology 79: 8339-8347.

Beasley DWC, Davis CT, Guzman H, Vanlandingham DL, Travassos da Rosa APA, Parsons RE, Higgs S, Tesh RB, Barrett ADT. 2003. Limited Evolution of West Nile virus has occurred during its southwesterly spread in the United States. Virology 309: 190-195.

Beasley DWC, Li L, Suderman MT, Barrett, ADT. 2002. Mouse neuroinvasive phenotype of West Nile virus strains varies depending upon virus genotype. Virology 296: 17-23.

Berthet FH, Zeller G, Drouet M, Rauzier J, Digoutte J, Deubel V. 1997. Extensive nucleotide changes and deletions within the envelope gene of Euro-African West Nile viruses. Journal of General Virology 78: 2293–2297.

Blackburn NK, Thompson DL, Jupp PG. 1987. Antigenic relationship of West Nile strains by tire ratios calculated from cross-neutralization test results. Epidemiology of Infectious Diseases 99: 551-557.

Blaney JE, Manipon GG, Murphy BR, Whitehead SS. 2003b. Temperature sensitive mutations in the genes encoding the NS1, NS2A, NS3, and NS5 nonstructural proteins of dengue virus type 4 restrict replication in the brains of mice. Archives of Virology 148: 999-1006.

Blaney JE, Manipon GG, Firestone CY, Johnson DH, Hanson CT, Murphy BR, Whitehead SS. 2003a. Mutations which enhance the replication of dengue virus type 4 and an antigenic chimeric dengue virus type 2/4 vaccine candidate in Vero cells. Vaccine 21: 4317-4327.

Blitvich BJ, Fernández-Salas I, Contreras-Cordero JF, Loroño-Pino MA, Marlenee NL, Díaz FJ, Beaty BJ. 2004. Phylogenetic analysis of West Nile virus, Nuevo Leon State, Mexico. Emerging Infectious Diseases 10: 1314-17.

Blitvich BJ, Fernandez-Salas I, Contreras-Cordero JF, Marlenee NL, Gonzalez-Rojas JI, Komar N, Gubler DJ, Calisher CH, Beaty BJ. 2003. Serologic evidence of West Nile virus infection in horses, Coahuila State, Mexico. Emerging Infectious Diseases 9: 853-856.

Bosio CF, Fulton RE, Salasek ML, Beaty BJ, Black WC. 2000. Quantitative trait loci that control vector competence for dengue-2 virus in the mosquito Aedes aegypti. Genetics 156: 687-698.

Brault AC, Langevin SA, Bowen RA, Panella NA, Biggerstaff BJ, Miller BR, Komar N. 2004. Differential virulence of West Nile strains for American crows. Emerging Infectious Diseases 10: 2161-2168.

Briese T, Jia XY, Huang C, Grady LJ, Lipkin WI. 1999. Identification of a Kunjin/West Nile-like flavivirus in brains of patients with New York encephalitis. Lancet 354: 1261–2.

Brinton MA. 2002. The molecular biology of West Nile virus: a new invader of the Western Hemisphere. Annual Reviews in Microbiology 56: 371-402.

Bunning ML, Bowen RA, Cropp CB, Sullivan KG, Davis BS, Komar N, Godsey MS, Baker D, Hettler DL, Holmes DA, Biggerstaff BJ, Mitchell CJ. 2002. Experimental infection of horses with West Nile virus. Emerging Infectious Diseases 8: 380-6.

Burt FJ, Grobbelaar AA, Leman PA, Anthony FS, Gibson GV, Swanepoel R. 2002. Phylogenetic relationships of southern African West Nile virus isolates. Emerging Infectious Diseases 8: 820–6.

CDC, Arbonet. http://www.cdc.gov/ncidod/dvbid/westnile/index.html. Accessed on August 24, 2005.

CDC. 2002. Provisional surveillance summary of the West Nile virus epidemic-United States, January-November 2002. Morbidity and Mortality Weekly Report 51: 1129-1133.

Chu PWG and Westaway EG. 1985. Replication Strategy of Kunjin Virus - Evidence for Recycling Role of Replicative Form Rna As Template in Semiconservative and Asymmetric Replication. Virology 140: 68-79.

Cologna R and Rico-Hesse R. 2003. American genotype structures decrease dengue virus output from human monocytes and dendritic cells. Journal of Virology 77: 3929-3938.

Darwish MA and Ibrahim AH. 1975. Prevalence of antibodies to arboviruses in Egypt. Results of a serologic survey among 1,113 university students. American Journal of Tropical Medicine and Hygeine 24: 981-985.

Davis CT, Beasley DCW, Guzman H, Raj P, D'Anton M, Novak RJ, Tesh RB, Barrett ADT. 2003. Genetic variation among temporally and geographically distinct West Nile virus isolates collected in the United State, 2001 and 2002. Emerging Infectious Disease 9: 1423-29.

Davis CT, Beasley DCW, Siirin M, Guzman H, Tesh RB, Barrett ADT. 2004. Emergence of attenuated West Nile virus isolates in Texas, 2003. Virology 330: 342-350.

Davis BS, Chang GJ, Cropp B, Roehrig JT, Martin DA, Mitchell CJ, Bowen R, Bunning ML. 2001. West Nile virus recombinant DNA vaccine protects mouse and horse from virus challenge and expresses in vitro a noninfectious recombinant antigen that can be used in enzyme-linked immunosorbent assays. Journal of Virology 75: 4040-4047.

De Madrid AT and Porterfield JS. 1974. The flaviviruses (group B arboviruses): a cross-neutralization study. Journal of General Virology 23: 91-96.

Duarte EA, Novella IS, Weaver SC, Domingo E, Wainhobson S, Clarke DK, Moya A, Elena SF, Delatorre JC, Holland JJ. 1994. Rna Virus Quasi-Species - Significance for Viral Disease and Epidemiology. Infectious Agents and Disease-Reviews Issues and Commentary 3: 201-214.

Dunster LM, Gibson CA, Stephenson JR, Minor PD, Barrett ADT. 1990. Attenuation of virulence of flaviviurses following passage in HeLa cells. Journal of General Virology 71: 601-607.

Dupuis AP, Marra P, Kramer LD. 2003. Serological evidence of West Nile virus transmission, Jamaica, West Indies. Emerging Infectious Diseases 9: 860-863.

Eastman PS, Blair CD. 1985. Temperature-sensitive mutants of Japanese encephalitis virus. Journal of Virology 55: 611-616.

Ebel GD, Dupuis AP, Ngo K, Nicholas D, Kauffman E, Jones SA, Bernard KA, Kramer LD. 2001. Partial genetic characterization of West Nile virus strains, New York State, 2000. Emerging Infectious Diseases 7: 650–653.

Ebel GD, Carricaburu JE, Young DS, Bernard KA, Kramer LD. 2004. Genetic and phenotypic variation of West Nile virus in New York, 2000-2003. American Journal of Tropical Medicine and Hygiene 71: 493-500.

Elizondo-Quiroga D, Davis CT, Fernandez-Salas I, Escobar-Lopez R, Elizondo-Quiroga A, Gonzalez-Rojas JI, Cordero JFC, Guzman H, Travassos da Rosa A, Blitvich BJ, Beaty BJ, Barrett ADT, Tesh RB. 2005. West Nile virus isolation in human and mosquitoes, Mexico. Emerging Infectious Diseases 11: 1449-1452.

Endy TP and Nisalak A. 2002. Japanese encephalitis virus: Ecology and epidemiology. Japanese Encephalitis and West Nile. Viruses 267: 11-48.

Estrada-Franco JG, Navarro-Lopez R, Beasley DWC, Coffey L, Carrara AS, Travassos da Rosa A, Tesh RB, Guzman H, Barrett ADT, Weaver SC. 2003. West Nile virus in

Mexico: evidence of widespread circulation since July 2002. Emerging Infectious Diseases 9: 1604-1607.

Fonseca DM, Keyghobadi N, Malcolm CA, Mehmet C, Schaffner F, Mogi M, Fleischer RC, Wilkerson RC. 2004. Emerging vectors in the Culex pipiens complex. Science 303: 1535-1538.

Granwehr BP, Lillibridge KM, Higgs S, Mason PW, Aronson JF, Campbell GA, Barrett, ADT. 2004. West Nile virus: where are we now? Lancet Infectious Diseases 4: 547-556.

Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, Holmes EC. 2004. Unifying the epidemiological and evolutionary dynamics of pathogens. Science 303: 327-332.

Hammam HM, Clarke DH, Price WH. 1965. Antigenic variation of West Nile virus in relation to geography. American Journal of Epidemiology 82: 40-55.

Hammam HM and Price WH. 1966. Further observations on geographic variation in the antigenic character of West Nile and Japanese B viruses. American Journal of Epidemiology 83: 113-122.

Hanley KA, Manlucu LR, Gilmore LE, Blaney JE, Hanson CT, Murphy BR, Whitehead SS. 2003. A trade-off in replication in mosquito versus mammalian systems conferred by a point mutation in the NS4B protein of dengue virus type 4. Virology 312: 222-232.

Hanley KA, Lee JJ, Blaney JE, Murphy BR, Whitehead SS. 2002. Paired charge-to-alanine mutagenesis of dengue virus type 4 NS5 generates mutants with temperature-sensitive, host range, and mouse attenuation phenotypes. Journal of Virology 76: 525-531.

Hayes CG. 2001. West Nile virus: Uganda, 1937, to New York City, 1999. Annals of the New York Academy of Science 951: 25-37.

Hayes CG. 1989. West Nile fever, In: Monath TP (ed.), The Arboviruses: Epidemiology and Ecology, CRC Press, Boca Raton: 59-88.

Hayes EB, Komar N, Nasci RS, Montgomery SP, O'Leary DR, Campbell GL. 2005b. Epidemiology and transmission dynamics of West Nile Virus disease. Emerging Infectious Diseases 11: 1167-1173.

Hayes EB, Sejvar JJ, Zaki SR, Lanciotti RS, Bode AV, Campbell GL. 2005a. Virology, pathology, and clinical manifestations of West Nile virus disease. Emerging Infectious Diseases 11: 1174-1179.

Heinz FX and Allison SL. 2003. Flavivirus structure and membrane fusion. Advances in Virus Research 59: 63-97.

Higgs S, Schneider BS, Vanlandingham DL, Klingler KA, Gould EA. 2005. Nonviremic transmission of West Nile virus. Proceedings of the National Academy of Sciences of the United States of America 102: 8871-8874.

Hoffmann E, Krauss S, Perez D, Webby R, Webster RG. 2002. Eight-plasmid system for rapid generation of influenza virus vaccines. Vaccine 20: 3165-3170.

Holland JJ. 1996. Evolving virus plagues - Commentary. Proceedings of the National Academy of Sciences of the United States of America 93: 545-546.

Hollingshead PG, Brawner TA, Fleming TP. 1983. St. Louis encephalitis virus temperature-sensitive mutants. I. Induction, isolation and preliminary characterization. Archives of Virology 75: 171-179.

Holmes EC and Twiddy SS. 2003. The origin, emergence and evolutionary genetics of dengue virus. Infectious Disease Genetic Evolution 3: 19-28.

Huang C, Slater BR, Rudd N, Parchuri R, Hull M, Dupuis M, Hindenberf A. 2002. First isolation of West Nile virus from a patient with encephalitis in the United States. Emerging Infectious Diseases 8: 1367-1371.

Huelsenbeck JP, and Ronquist FR. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754-55.

Iwamoto M, Jernigan DB, Guasch A, Trepka MJ, Blackmore CG, Hellinger WC, Pham, SM, Zaki S, Lanciotti RS, Lance-Parker SE, Diaz-Granados CA, Winquist AG, Perlino, CA, Wiersma S, Hillyer KL, Goodman JL, Marfin AA, Chamberland ME, Petersen LR. 2003. Transmission of West Nile virus from an organ donor to four transplant recipients. New England Journal of Medicine 348: 2196-2203.

Jia XY, Briese T, Jordan I, Rambaut A, Chi HC, Mackenzie JS, Hall RA, Scherret J, Lipkin WI. 1999. Genetic analysis of West Nile New York 1999 encephalitis virus. Lancet 354: 1971-1972.

Kapoor M, Zhang LW, Ramachandra M, Kusukawa J, Ebner KE, Padmanabhan R. 1995. Association Between Ns3 and Ns5 Proteins of Dengue Virus Type-2 in the Putative Rna Replicase Is Linked to Differential Phosphorylation of Ns5. Journal of Biological Chemistry 270: 19100-19106.

Khromykh AA and Westaway EG. 1996. RNA binding properties of core protein of the flavivirus Kunjin. Archives of Virology 141: 685-699.

Khromykh AA, Sedlak PL, Westaway EG. 1999. trans-complementation analysis of the flavivirus Kunjin ns5 gene reveals an essential role for translation of its N-terminal half in RNA replication. Journal of Virology 73: 9247-9255.

Klenk K, Snow J, Morgan K, Bowen R, Stephens M, Foster F, Gordy P, Beckett S, Komar N, Gubler D, Bunning M. 2004. Alligators as West Nile virus amplifiers. Emerging Infectious Disease 10: 2150-2155.

Komar N. 2001. West Nile virus surveillance using sentinel birds. Annals of the New York Academey of Science 951: 58-73.

Komar N, Lanciotti R, Bowen R, Langevin S, Bunning M. 2002. Detection of West Nile virus in oral and cloacal swabs collected from bird carcasses. Emerging Infectious Disease 8: 741-742.

Komar N, Langevin S, Hinten S, Nemeth N, Edwards E, Hettler D, Davis B, Bowen R, Bunning M. 2003. Experimental infection of North American birds with the New York 1999 strain of West Nile virus. Emerging Infectious Disease 9: 311-322.

Lai CJ and Monath TP. 2003. Chimeric flaviviruses: novel vaccines against dengue fever, tick-borne encephalitis, and Japanese encephalitis. Advances in Virus Research 61: 469-509.

Lanciotti RS, Roehrig JT, Deubel V, Smith J, Parker M, Steele K, Volpe KE, Crabtree MB, Scherret JH, Hall RA, MacKenzie JS, Cropp CB, Panigrahy B, Ostlund E, Schmitt B, Malkinson M, Banet C, Weissman J, Komar N, Savage HM, Stone W, McNamara T, Gubler DJ. 1999. Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern U.S. Science 286: 2333–2337.

Lanciotti RS, Ebel GD, Deubel V, Kerst AJ, Murri S, Meyer R, Bowen M, McKinney N, Morrill WE, Crabtree MB, Kramer LD, Roehrig JT. 2002. Complete genome sequences and phylogenetic analysis of West Nile virus strains isolated from the United States, Europe, and the Middle East. Virology 298: 96-105.

Ledger TN, Sil BK, Dunster LM, Stephenson JR, Minor PD, Barrett ADT. 1992. Yellow fever 17DD vaccine virus is temperature sensitive when grown in mosquito C6-36 cells. Vaccine 10: 652-654.

Leitmeyer KC, Vaughn DW, Watts DM, Salas R, de Chacon IV, Ramos C, and Rico-Hesse R. 1999. Dengue virus structural differences that correlate with pathogenesis. Journal of Virology 73: 4738-4747.

Lillibridge KM, Parsons R, Randle Y, Travassos da Rosa APA, Guzman H, Siirin M, Wuithiranyagool T, Hailey C, Beasley DW, Higgs S, Pascual R, Meyer T, Barrett ADT, Tesh RB. 2004. The 2002 introduction of West Nile virus into Harris County, TX, an area historically endemic for St. Louis encephalitis. American Journal of Tropical Medicine and Hygiene 70: 676-681.

Lindenbach BD and Rice CM. 1999. Genetic interaction of flavivirus nonstructural proteins NS1 and NS4A as a determinant of replicase function. Journal of Virology 73: 4611-4621.

Lindenbach BD and Rice CM. 2003. Molecular biology of flaviviruses. Advances in Virus Research 59: 23-61.

Ma L, Jones CT, Groesch TD, Kuhn RJ, Post CB. 2004. Solution structure of dengue virus capsid protein reveals another fold. Proceedings of the National Academy of Science, USA 101: 3414-3419.

Mandl CW, Aberle JH, Aberle SW, Holzmann H, Allison SL, Heinz FX. 1998. In vitro-synthesized infectious RNA as an attenuated live vaccine in a flavivirus model. Nature Medicine 4: 1438-1440.

Markoff L. 2003. 5'- and 3'-noncoding regions in flavivirus RNA. Advance in Virus Research 59: 177-228.

Miller BR, Nasci RS, Godsey MS, Savage HM, Lutwama JJ, Lanciotti RS, Peters CJ. 2000. First field evidence for natural vertical transmission of West Nile virus in Culex univittatus complex mosquitoes from Rift Valley province, Kenya. American Journal of Tropical Medicine and Hygiene 62: 240-246.

Minke JM, Siger L, Karaca K, Austgen L, Gordy P, Bowen R, Renshaw RW, Loosmore S, Audonnet JC, Nordgren B. 2004. Recombinant canarypoxvirus vaccine carrying the prM/E genes of West Nile virus protects horses against a West Nile virus-mosquito challenge. Archives of Virology, Supplement: 221-230.

Modis Y, Ogata S, Clements D, Harrison SC. 2003. A ligand-binding pocket in the dengue virus envelope glycoprotein. Proceedings of the National Academy Science USA. 100: 6986-6991.

Monath TP. 2002. Editorial: Jennerian vaccination against West Nile virus. American Journal of Tropical Medicine Hygiene 66: 113-114.

Morvan J, Fontenille D, Lepers JP, Coulanges P. 1990. Serological surveillance of West Nile virus infection in a village of the Central Highlands in Madagascar. Annals of Social Belgian Tropical Medicine 70: 55-63.

Mukhopadhyay S, Kim BS, Chipman PR, Rossmann MG, Kuhn RJ. 2003. Structure of West Nile virus. Science 302: 248-249.

Munoz-Jordan JL, Sanchez-Burgos GG, Laurent-Rolle M, Garcia-Sastre A. 2003. Inhibition of interferon signaling by dengue virus. Proceedings of the National Academy of Science, USA 100:14333-14338.

Murgue B, Zeller H, Duebel V. 2002. The ecology and epidemiology of West Nile virus in Africa, Europe, and Asia. Current Topics in Microbiology Immunology 267: 195-221.

Muylaert IR, Galler R, Rice CM. 1997. Genetic analysis of the yellow fever virus NS1 protein: Identification of a temperature-sensitive mutation which blocks RNA accumulation. Journal of Virology 71: 291-298.

Nalca A, Fellows PF, Whitehouse CA. 2003. Vaccines and animal models for arboviral encephalitides. Antiviral Res. 60: 153-174.

Ng T, Hathaway D, Jennings N, Champ D, Chiang YW, Chu HJ. 2003. Equine vaccine for West Nile virus. Developmental Biology (Baseline) 114: 221-227.

Ni H, Chang GJ, Xie H, Trent DW, Barrett ADT. 1995. Molecular basis of attenuation of neurovirulence of wild-type Japanese encephalitis virus strain SA14. Journal of General Virology 76: 409-413.

Novella IS, Clarke DK, Quer J, Duarte EA, Lee CH, Weaver SC, Elena SF, Moya A, Domingo E, Holland JJ. 1995. Extreme fitness differences in mammalian and insect hosts after continuous replication of vesicular stomatitis virus in sandfly cells. Journal of Virology 69: 6805-6809.

Omilabu SA, Olaleye OD, Aina Y, Fagbami AH. 1990. West Nile Complement Fixing antibodies in Nigerian domestic animals and humans. Journal of Hygiene and Epidemiological Microbiology and Immunology 34: 357-363.

Pealer LN, Marfin AA, Petersen LR, Lanciotti RS, Page PL, Stramer SL, Stobierski MG, Signs K, Newman B, Kapoor H, Goodman JL, Chamberland ME. 2003. Transmission of West Nile virus through blood transfusion in the United States in 2002. New England Journal of Medicine 349: 1236-1245.

Peiris JSM and Amerasinghe FP. 1994. West Nile fever. In: Beran GW and Steele JH, editors. Handbook of zoonoses. 2nd ed. Section B: Viral. Boca Raton (FL): CRC Press: 139–148.

Pletnev AG, Claire MS, Elkins R, Speicher J, Murphy BR, Chanock RM. 2003. Molecularly engineered live-attenuated chimeric West Nile/dengue virus vaccines protect rhesus monkeys from West Nile virus. Virology 314: 190-195.

Price WH and O'Leary W. 1967. Geographic variation in the antigenic character of West Nile virus. American Journal of Epidemiology 85: 84–86.

Proutski V, Gritsun TS, Gould EA, Holmes EC. 1999. Biological consequences of deletions within the 3'-untranslated region of flaviviruses may be due to rearrangements of RNA secondary structure. Virus Research 64: 107-123.

Puri B, Nelson WM, Henchal EA, Hoke CH, Eckels KH, Dubois DR, Porter KR, Hayes CG. 1997. Molecular analysis of dengue virus attenuation after serial passage in primary dog kidney cells. Journal of General Virology 78: 2287-2291.

Quirin R, Salas M, Zientara S, Zeller H, Labie J, Murri S, Lefrancois T, Petitclerc M, Martinez D. 2004. West Nile virus, Guadeloupe. Emerging Infectious Diseases 10: 706-708.

Rappole JH, Derrickson SR, Hubalek Z. 2000. Migratory birds and spread of West Nile virus in the Western Hemisphere. Emerging Infectious Disease 6: 319–328.

Rappole JH and Hubalek Z. 2003. Migratory birds and West Nile virus. Journal of Applied Microbiology 94: 47–58.

Reisen WK. 2003. Epidemiology of St. Louis encephalitis virus. Advances in Virus Research 61: 139-183.

Rico-Hesse R, Harrison LM, Nisalak A, Vaughn DW, Kalayanarooj S, Green S, Rothman AL, Ennis FA. 1997. Molecular evolution of dengue type 2 virus in Thailand. American Journal of Tropical Medicine and Hygiene 58: 96-101.

Ronquist F and Huelsenbeck JP. 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572-1574.

Rossi SL, Zhao Q, O'Donnell VK, Mason PW. 2005. Adaptation of West Nile virus replicons to cells in culture and use of replicon-bearing cells to probe antiviral action. Virology 331: 457-470.

Scherret JM, Poidinger M, MacKenzie JS, Broom AK, Deubel V, Lipkin WI, Briese T, Gould EA and Hall RA. 2001. Studies on the relationships between West Nile and Kunjin viruses. Emerging Infectious Diseases 11: 697-705.

Shirato K, Miyoshi H, Goto A, Ako Y, Ueki T, Kariwa H, Takashima I. 2004. Viral envelope protein glycosylation is a molecular determinant of the neuroinvasiveness of the New York strain of West Nile virus. Journal of General Virology 85: 3637-3645.

Smithburn KC. 1942. Differentiation of the West Nile virus from the viruses of St. Loius and Japanese B encephalitis. Journal of Immunology 44: 25-31.

Smithburn, KC, Hughes, TP, Burke AW, Paul JH. 1940. A neurotropic virus isolated from the blood of a native of Uganda. American Journal of Tropical Medicine 20: 471-492.

Solomon T, Fisher AF, Beasley DW, Mandava P, Granwehr BP, Langsjoen H, Travassos Da Rosa AP, Barrett AD, Tesh RB. 2003. Natural and nosocomial infection in a patient with West Nile encephalitis and extrapyramidal movement disorders. Clinical Infectious Diseases 36: 140-145.

Solomon T, Vaughn DW. 2002. Pathogenesis and clinical features of Japanese encephalitis and West Nile virus infections. Current Topics in Microbiology and Immunology 267: 171-194.

Southam CM and Moore AE. 1954. Induced virus infections in man by the Egypt isolates of West Nile virus. American Journal of Tropical Medicine and Hygiene 3: 19-50.

Steel KE, Linn MF, Schoepp RJ, Komar N, Geisbert TW, Manduca RM. 2000. Pathology of fatal West Nile virus infections in native and exotic birds during the 1999 outbreak in New York City, New York. Veterinary Pathology 37: 208–224.

Stobierski MG, Signs K, Newman B, Kapoor H, Goodman JL, Chamberland ME; West Nile Virus Transmission Investigation Team. 2003. Transmission of West Nile virus through blood transfusion in the United States in 2002. New England Journal of Medicine 349: 1236-1245.

Swofford DL. 2002. PAUP: Phylogenetic analysis using parsimony (and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Taylor RM, Work TH, Hurlbut HS, Rizk F. 1956. A study of the ecology of West Nile virus in Egypt. American Journal of Tropical Medicine 5: 579-620.

Tesh RB, Parsons R, Siirin M, Randle Y, Sargent C, Guzman H, Higgs S. 2004. Year-round West Nile virus activity, Gulf Coast Region, Texas and Louisiana. Emerging Infectious Disease 10: 1649-1652.

Tesh RB, Travassos da Rosa APA, Guzman H, Araujo TP, Xiao SY. 2001. Immunization with heterologous flaviviruses protects against fatal West Nile encephalitis. Emerging Infectious Diseases 8: 245-251.

Tilgner M, Deas TS, Shi PY. 2005. The flavivirus-conserved penta-nucleotide in the 3' stem-loop of the West Nile virus genome requires a specific sequence and structure for RNA synthesis, butnot for viral translation. Virology 3312: 375-386.

Tonry JH, Brown CB, Cropp CB, Co JK, Bennett SN, Nerurkar VR, Kuberski T, Gubler DJ. 2005. West Nile virus detection in urine. Emerging Infectious Diseases 11: 1294-1296.

Usuku S, Castillo L, Sugimoto C, Noguchi Y, Yogo Y, Kobayashi N. 2001. Phylogenetic analysis of dengue-3 viruses prevalent in Guatemala during1996-1998. Archives of Virology 146: 1381-1390.

Wallner G, Mandl CW, Ecker M, Holzmann H, Stiasny K, Kunz C, Heinz FX. 1996. Characterization and complete genome sequences of high- and low- virulence variants of tick-borne encephalitis virus. Journal of General Virology 77: 1035-1042.

Wang T, Town T, Alexopoulou L, Anderson JF, Fikrig E, Flavell RA. 2004. Toll-like receptor 3 mediates West Nile virus entry into the brain causing lethal encephalitis. Nature Medicine 10: 1366-1373.

Watts DM, Porter KR, Putvatana P, Vasquez B, Calampa C, Hayes CG, Halstead SB. 1999. Failure of secondary infection with American genotype dengue 2 to cause dengue haemorrhagic fever. Lancet 354: 1431-1434.

Weaver SC, Brault AC, Kang W, Holland JJ. 1999. Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. Journal of Virology 73: 4316-4326.

Westaway EG, Khromykh AA, Kenney MT, Mackenzie JM, Jones MK. 1997. Proteins C and NS4B of the flavivirus Kunjin translocate independently into the nucleus. Virology 234: 31-41.

Wittke V, Robb TE, Thu HM, Nisalak A, Nimmannitya S, Kalayanrooj S, Vaughn DW, Endy TP, Holmes EC, Aaskov JG. 2002. Extinction and rapid emergence of strains of dengue 3 virus during an interepidemic period. Virology 301: 148-156.

Xiao SY, Guzman H, Zhang H, Travassos da Rosa AP, Tesh RB. 2001. West Nile virus infection in the golden hamster (Mesocricetus auratus): a model for West Nile encephalitis. Emerging Infectious Diseases 7: 714-721.

VITA

I, Charles Todd Davis, was born on April 1st, 1975 in Birmingham, AL where my parents, James and Mary Todd Davis now live. My wife, Lauren Davis, and I currently live at 601 8th St. Galveston, TX 77550. After receiving a B.S. in Biology from the University of the South with an interest in zoonotic diseases, I accepted an internship with the Bronx Zoo's St. Catherine Island Wildlife Survival Center. Working with veterinarians responsible for the welfare of endangered animals at the breeding facility, I was introduced to issues concerning the transmission of diseases and the diagnostic techniques and systems designed to monitor and manage these problems. After completion of this internship, my interests in laboratory disease diagnosis became a priority. This desire led to an internship with the Center of Medical Investigation in Santiago, Chile as a research assistant to epidemiologists studying bacterial, viral, and parasitic diseases. Field research pertaining to the control of the bacterium *Bartonella henselae*, the etiological agent of Cat Scratch Disease, and training in an infectious disease lab proved an excellent opportunity to acquire many diagnostic skills. Besides becoming more proficient in Spanish, the work reinforced my knowledge of infectious diseases and heightened my awareness of current biomedical research. Ultimately, this led to my matriculation into Tulane's School of Public Health and Tropical Medicine where I received a Master of Science in Public Health while conducting research in Guatemala aimed at assessing the success of a community-based dengue fever prevention program. The combination of an MSPH in Tropical Medicine and research experience abroad then encouraged me to pursue a degree at the doctoral level at the University of Texas Medical Branch (UTMB). As a UTMB graduate student, I have been fortunate to work in the laboratory of Dr. Alan Barrett and have taken on a dissertation project studying the molecular epidemiology and pathogenesis of West Nile virus in North America. My dissertation research has focused on the evolution of West Nile virus since its introduction into the Western Hemisphere and, in particular, how genetic changes to this virus have resulted in phenotypic variation among virus isolates.

Education

Ph. D., December 2005, University of Texas Medical Branch, Galveston, TX

M.S.P.H., January 2001, Tulane University School of Public Health and Tropical Medicine, New Orleans, LA

B.S. May 1997, University of the South. Sewanee, TN

Publications

Beasley DCW, Davis CT, Guzman H, Vanlandingham DL, Travassos da Rosa APA, Parsons RE, Higgs S, Tesh RB, Barrett ADT. 2003. Limited evolution of West Nile virus during its southwesterly spread in the United States. Virology 309:190-195.

Davis CT, Beasley DCW, Guzman H, Raj P, D'Anton M, Novak RJ, Unasch TR, Tesh RB, Barrett ADT. 2003. Genetic variation among temporally and geographically distinct West Nile virus isolates collected in the Unites States, 2001 and 2002. Emerging Infectious Diseases 9:1423-1429.

Beasley DWC, Davis CT, Whiteman M, Granwehr B, Kinney RM and Barrett ADT. 2004. Molecular determinants of virulence of West Nile virus in North America. In "Emergence and control of zoonotic viral encephalitides." Edited by CH Calisher and DE Griffin. Archives of Virology supplement 18, pp. 35-41.

Moncayo AC, Fernandez Z, Diallo M, Sall A, Hartman S, Davis CT, Coffey L, Ortiz D, Mathiot CC, Tesh RB, Weaver SC. 2004. Emergence of Epidemic Dengue Through the Adaptation of Sylvatic Progenitor Viruses to Anthropophilic Mosquitoes. Emerging Infectious Diseases 10(10): 1790-1796.

Davis CT, Beasley DWC, Guzman H, Siirin M, Parsons RE, Tesh RB, Barrett ADT. 2004. Emergence of attenuated West Nile virus variants in Texas, 2003. Virology. 330: 342-350.

Granwehr BP, Li Li, Davis CT, Beasley DCW, Barrett ADT. 2004. Characterization of a human isolate of West Nile Virus from the Gulf Coast of Texas. J Clin Microbiol. 42(11): 5375-7.

Beasley DWC, Davis CT, Estrada-Franco J, Navarro-Lopez R, Campomanes-Cortes A, Tesh RB, Weaver SC, Barrett ADT. 2004. Complete genome sequence of a West Nile virus isolate from Mexico and identification of an attenuating mutation. Emerging Infectious Disease. 10(12): 2221-2224.

Beasley DWC, Davis CT, Whiteman M, Barrett ADT. 2004. Molecular determinants of virulence variation between West Nile virus strains. In "Arbovirus Research in Australia-Vol. 9." Edited by PA Ryan, JG Aaskov, TD St. George, and PER Dale. pp. 24-27.

Elizondo-Quiroga D, Davis CT, Fernandez-Salas I, Escobar-Lopez R, Elizondo-Quiroga A, Gonzalez-Rojas JI, Cordero JFC, Guzman H, Travassos da Rosa A, Blitvich BJ, Beaty BJ, Barrett ADT, Tesh RB. 2005. West Nile virus isolation in human and mosquitoes, Mexico. Emerging Infectious Diseases 11(9): 1449-1452.

Davis CT, Ebel GD, Lanciotti RS, Brault AC, Guzman H, Siirin M, Lambert A, Parsons RE, Beasley DWC, Novak RJ, Elizondo-Quiroga D, Green EN, Young DS, Stark LM, Drebot MA, Artsob H, Tesh RB, Kramer LD, Barrett ADT. Phylogenetic Analysis of North American West Nile virus Isolates, 2001-2004: Evidence for the Emergence of a Dominant Genotype. Virology. In press.