DISSERTATION

# INVESTIGATIONS OF THE EVOLUTIONARY, EPIDEMIC, AND MAINTENANCE POTENTIAL OF LA CROSSE VIRUS

Submitted by

Sara M. Reese

Department of Microbiology, Immunology, and Pathology

In partial fulfillment of the requirements

For the Degree of Doctorate of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2008

COLORADO STATE UNIVERSITY

April 7th, 2008

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY SARA M. REESE ENTITLED INVESTIGATIONS OF THE EVOLUTIONARY, EPIDEMIC, AND MAINTENANCE POTENTIAL OF LA CROSSE VIRUS BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

COMMITTEE ON GRADUATE WORK

Carol D. Blair

Carol D. Blair

William C. Black IV

Chester G. Moore

Richard A. Bowen

**Advisor** Barry J. Beaty

**Department Head** Carol D. Blair

## ABSTRACT OF DISSERTATION

## INVESTIGATIONS OF THE EVOLUTIONARY, EPIDEMIC, AND MAINTENANCE POTENTIAL OF LA CROSSE VIRUS

Arthropod-borne viruses (arboviruses) are resurging and emerging worldwide, and La Crosse virus (LACV) is a prototypical emergent virus in the United States. In this dissertation, the evolutionary, epidemic, and maintenance potential of LACV is investigated. In laboratory and field studies, LACV has shown significant evolutionary and epidemic potential through point mutations and segment reassortment. Through the use of sensitive molecular epidemiological techniques, significant genetic variation was observed in LACV RNA amplified from field-infected *Aedes triseriatus* mosquitoes, suggesting the potential for frequent segment reassortment of LACV in nature. Maximum parsimony phylogenetic analysis and linkage disequilibrium analysis revealed that 25-38.6% of the mosquito samples contained reassortant viruses. The geographical, environmental and temporal factors that condition the genetic structure of LACV were also investigated. The analysis revealed that there are no physical barriers to viral gene flow in the study site, indicating that the more or less virulent LACV strains could traffic and be transmitted throughout the entire 15,360 $km^2$ study range (southeastern Wisconsin, southwestern Minnesota and northeastern Iowa). Although there were no barriers to viral gene flow and no isolation by distance, a significant temporal association with viral genotype was revealed.

The maintenance of LACV in nature is not well understood. Mathematical models have revealed that field infection rates are well below those required to maintain the virus in nature. However, the mathematical models have not considered the possibilities of

stably-infected *Ae. triseriatus* mosquitoes or a LACV induced mating advantage for infected females. Super-infected *Ae. triseriatus* mosquitoes in nature were identified in southeastern Wisconsin and southwestern Minnesota in these studies (0.011% prevalence rate) suggesting that LACV could be maintained in nature through a stabilized infection in a small number of females.

LACV maintenance in nature may also be assisted by a mating advantage for *Ae. triseriatus* females infected with LACV. In this study, LACV transovarially-infected female mosquitoes become inseminated faster than uninfected mosquitoes, and this could increase the chance for transovarial transmission as well as venereal transmission of the virus. The evolutionary and maintenance potential of LACV was investigated in this dissertation and the results provide insight into the determinants of arbovirus emergence and epidemic potential in nature.

Sara M. Reese
Microbiology, Immunology and Pathology Department
Colorado State University
Fort Collins, CO 80523
Spring 2008

## ACKNOWLEDGEMENTS

First, I would like to thank my advisor, Barry Beaty, for his continuous support throughout my PhD program. Without his guidance, I would never have learned how to think like a scientist, speak like a virologist, and most importantly, learn the importance of bio-politics. I appreciate that he did not give up on me when I most likely provided him ample opportunities to do so. Working with Dr. Beaty has been a rewarding, memorable and challenging experience that I will forever be grateful.

I would also like to thank the rest of my committee that has been extremely supportive and helpful throughout my program. Carol Blair was always there to calm my nerves and answer my multitude of questions. I also thoroughly enjoyed spending hours in Bill Black's office analyzing data. Without my committee's support and faith in me, this work would not have been completed.

My family and friends have also been extremely supportive throughout my PhD program. I would especially like to thank my husband, Jeremy, for his never ending support and faith in me. My mother also deserves special mention for her encouragement when I doubted myself. My sister, Kelly, was always there ready with something to make me laugh after a long day in the lab. My personal motivational speaker, Amy, was also imperative to my program because without her advice and confidence building lectures, I would have not been able to survive the challenges I faced.

Finally, I would like to acknowledge the Fellowship Training Program that is funded by the Centers for Disease Control and Prevention. This grant funded me through my program.

# DEDICATION

This work is dedicated to my father, J. Lee Kimbrell, who would have been thrilled with this great accomplishment.

## TABLE OF CONTENTS


I. LITERATURE REVIEW ..... 1
A. INTRODUCTION ..... 2
1. The global resurgence of arbovirus diseases ..... 2
2. Emergence and resurgence of selected arboviruses ..... 4
B. ARTHROPOD-BORNE VIRUSES ..... 7
1. Arthropod-borne (Arbo) virus transmission cycles ..... 7
2. Vector competence ..... 8
3. Vectorial capacity ..... 9
4. Mechanical, horizontal and vertical transmission of arboviruses ..... 10
5. Virus adaptation in an arthropod host ..... 11
C. *BUNYAVIRIDAE* ..... 13
1. *Bunyaviridae* family overview ..... 13
2. *Orthobunyavirus* genus ..... 14
3. Molecular biology and gene structure-function relationships of *Bunyaviruses* 14
D. LA CROSSE VIRUS ..... 17
1. La Crosse virus epidemiology and clinical manifestations ..... 17
2. La Crosse virus transmission cycle ..... 18
3. La Crosse virus pathogenesis in vertebrate hosts ..... 19
4. *Aedes triseriatus*-LACV interactions ..... 20
E. QUASISPECIES – A MODEL FOR ARBOVIRUS TRANSMISSION CYCLES ..... 23
F. EVOLUTIONARY POTENTIAL OF THE *BUNYAVIRIDAE* FAMILY ..... 24
1. Mechanism for the generation of genetic diversity ..... 24
2. Evolutionary potential of La Crosse virus ..... 25
3. Segment reassortment of other *Bunyaviruses* ..... 29
G. MOLECULAR EPIDEMIOLOGY: DEFINING ARBOVIRUS DETERMINANTS OF EMERGENCE AND RESURGENCE ..... 32
1. The basis of molecular epidemiology ..... 32
2. St. Louis encephalitis (SLE) virus ..... 32
3. Dengue virus ..... 33
4. Molecular epidemiology of *Bunyaviruses* ..... 34
H. RATIONALE AND OUTLINE OF DISSERTATION ..... 38

II. POTENTIAL FOR SEGMENT REASSORTMENT OF LA CROSSE VIRUS IN NATURE ..... 40
A. INTRODUCTION ..... 41
B. MATERIALS AND METHODS ..... 44
1. Egg collection ..... 44
2. LACV analytes ..... 44
3. Immunofluorescence assay (IFA) ..... 46
4. RNA purification amplification by reverse transcription-PCR ..... 46
5. Amplicon cloning and sequencing ..... 47
6. Haplotype determination ..... 48
7. Statistical analyses ..... 48
C. RESULTS ..... 50

1. LACV infected mosquitoes analyzed ..... 50
2. Sequence polymorphism and nucleotide substitution ..... 50
3. Haplotype determination ..... 53
4. Phylogenetic analysis ..... 57
5. Linkage disequilibrium ..... 57
6. Occurrence of quasispecies ..... 65
D. DISCUSSION ..... 66

III. GENOTYPIC VARIABILITY OF LA CROSSE VIRUS: INVESTIGATIONS OF GEOGRAPHICAL (SPATIAL), ENVIRONMENTAL, AND TEMPORAL FACTORS THAT CONDITION THE GENETIC STRUCTURE OF THE VIRUS IN AN ENDEMIC AREA ..... 71
A. INTRODUCTION ..... 72
B. MATERIALS AND METHODS ..... 76
1. Egg collection ..... 76
2. Immunofluorescence assay ..... 77
3. RNA purification amplification by reverse transcription-PCR ..... 77
4. Amplicon cloning and sequencing ..... 79
5. AMOVA analysis ..... 79
6. Determination of land cover types ..... 81
7. Distance Matrix and Phylogenetic Analysis ..... 81
8. Haplotype distribution ..... 81
C. RESULTS ..... 82
1. Study population ..... 82
2. Haplotype determination ..... 82
3. AMOVA for determination of isolation by distance ..... 82
4. AMOVA for determination of physical barriers to viral gene flow ..... 85
5. AMOVA for determination of ecological determinants ..... 86
6. AMOVA for determination of temporal effects of LACV haplotypes ..... 89
7. Chi-square test for independence of virus haplotype and collection year ..... 92
8. Detection of quasispecies in LACV-infected mosquitoes ..... 92
D. DISCUSSION ..... 97

IV. IDENTIFICATION AND CHARACTERIZATION OF LA CROSSE VIRUS ISOLATED FROM SUPER-INFECTED *AEDES TRISERATUS* MOSQUITOES FROM THE FIELD ..... 104
A. INTRODUCTION ..... 105
B. METHODS AND MATERIALS ..... 111
1. Egg collection ..... 111
2. Immunofluorescence assay ..... 111
3. Virus isolation ..... 114
4. $TCID_{50}$ ..... 114
5. Plaque purification ..... 114
6. RNA purification from virus isolates and amplification by reverse transcription-PCR ..... 115
7. LACV RNA purification and amplification by reverse transcription-PCR from infected mosquitoes ..... 117

8. Sequencing ..... 117
9. Maximum parsimony analysis ..... 117
10. Nucleotide sequence divergence between populations ..... 118
C. RESULTS ..... 118
1. LACV-super-infected mosquito prevalence in the field ..... 118
2. LACV titer of super-infected mosquitoes ..... 123
3. Phylogenetic analysis of virus isolates from SI-mosquitoes ..... 124
4. Phylogenetic analysis of MAS from SI-mosquitoes and NI-mosquitoes ..... 128
5. Sequence polymorphism, nucleotide substitution and nucleotide diversity ..... 132
6. Nucleotide diversity ..... 134
7. Nucleotide sequence divergence ..... 135
D. DISCUSSION ..... 139

V. COMPARISON OF INSEMINATION RATES OF LA CROSSE VIRUS TRANSOVARIALLY INFECTED AND UNINFECTED *AEDES TRISERIATUS* MOSQUITOES ..... 148
A. INTRODUCTION ..... 149
B. METHODS AND MATERIALS ..... 152
1. Laboratory mosquitoes ..... 152
2. Field mosquitoes ..... 152
3. Immunofluorescence assay ..... 152
4. Bloodmeals ..... 153
5. Mosquito mating ..... 153
6. Determination of genome equivalents through quantitative reverse transcription PCR (qRT-PCR) ..... 154
7. Cuticular hydrocarbon analysis ..... 155
8. Female accessory sex gland dissection and staining ..... 156
9. Statistics ..... 156
C. RESULTS ..... 157
1. Comparison of insemination rates of laboratory colonized non-bloodfed LACV- and LACV+ mosquitoes ..... 157
2. Comparison of insemination rates of laboratory colonized bloodfed LACV- and LACV+ mosquitoes ..... 157
3. Comparison of insemination rates of non-bloodfed LACV- and LACV+ field collected mosquitoes ..... 158
4. Comparison of insemination rates of LACV- and LACV+ field-collected and bloodfed mosquitoes ..... 160
5. Association of genome equivalents and insemination ..... 160
6. Cuticular Hydrocarbon Analysis ..... 163
7. Female accessory sex gland ..... 165
D. DISCUSSION ..... 166

VI. Summary and Discussion ..... 172
I. SUMMARY AND DISCUSSION ..... 173

REFERENCES ..... 204

## TABLE OF FIGURES

Figure 2.1: Mosquito collection sites in Minnesota, Wisconsin, and Iowa ..... 45
Figure 2.2: Nucleotide polymorphism (theta) of the LACV L, M and S segment sequences amplified from field-infected mosquitoes ..... 52
Figure 2.3: LACV L, M, and S segment haplotype determination based on the original approach ..... 55
Figure 2.4: LACV L, M, and S segment haplotype determination based on the conservative approach ..... 56
Figure 2.5: Maximum parsimony phylogenetic analysis of the L segment of LACV RNA amplified from field collected mosquitoes ..... 58
Figure 2.6: Maximum parsimony phylogenetic analysis of the M segment of LACV RNA amplified from field collected mosquitoes ..... 59
Figure 2.7: Maximum parsimony phylogenetic analysis of the S segment of LACV RNA amplified from field collected mosquitoes ..... 60
Figure 2.8: Evidence for multiple virus haplotypes in a single mosquito ..... 65

Figure 3.1: Mosquito collection sites in Minnesota, Wisconsin and Iowa. ..... 78
Figure 3.2: Haplotype distribution for the S and M segment sequences of LACV amplified from mosquitoes collected in 2000, 2004, and 2005 ..... 84
Figure 3.3: S segment maximum parsimony tree reveals haplotypes segregating into clades by collection year. ..... 93
Figure 3.4: M segment maximum parsimony tree revealed haplotypes cluster into clades by collection year ..... 94
Figure 3.5: The frequency of haplotypes is significantly associated with the collection year. ..... 95
Figure 3.6: Evidence for multiple virus haplotypes in a single mosquito ..... 96

Figure 4.1: Mosquito collection sites in Minnesota, Wisconsin and Iowa. ..... 112
Figure 4.2: IFA images from LACV-infected field mosquitoes ..... 113
Figure 4.3: The distribution of LACV-infected mosquitoes collected in 2007. ..... 122
Figure 4.4: Maximum parsimony phylogenetic analysis of the entire S segment of LACV isolates ..... 125
Figure 4.5: Maximum parsimony phylogenetic analysis of the entire M segment of LACV isolates ..... 126
Figure 4.6: Maximum parsimony phylogenetic analysis of the entire L segment of LACV isolates ..... 127
Figure 4.7: Maximum parsimony phylogenetic analysis of a portion of the S segment RNA of MAS from normally-infected and super-infected field collected mosquitoes ..... 129
Figure 4.8: Maximum parsimony phylogenetic analysis of a portion of the M segment RNA of MAS from normally-infected and super-infected field collected mosquitoes ..... 130

Figure 4.9: Maximum parsimony phylogenetic analysis of a portion of the L segment RNA of MAS from normally-infected and super-infected field collected mosquitoes ..... 131
Figure 4.10: Nucleotide polymorphism (theta) of the LACV S, M and L ..... 133
Figure 4.11: Nucleotide sequence divergence ($D_{xy}$) of the LACV S, M and L ..... 138

Figure 5.1: Female accessory sex gland of *Ae. triseriatus* ..... 156
Figure 5.2: LACV-TOT+ laboratory colonized females were inseminated more rapidly compared to uninfected mosquitoes without a bloodmeal ..... 158
Figure 5.3: LACV-TOT+ laboratory colonized females were inseminated more rapidly than the uninfected mosquitoes following a bloodmeal on day 3. ..... 159
Figure 5.4: Insemination rates were greater for LACV TOT+ field collected mosquitoes than for noninfected mosquitoes without and with a bloodmeal ..... 161
Figure 5.5: LACV titer is not correlated with insemination efficiency ..... 162
Figure 5.6: Chromatogram of *Aedes triseriatus* mosquitoes one day post-bloodmeal ..... 164
Figure 5.7: LACV infection of the female accessory sex gland ..... 165

# TABLE OF TABLES

Table 2.1: LACV isolates analyzed in the segment reassortant analysis ..... 46
Table 2.2: Field collected *Ae. triseriatus* mosquitoes from Minnesota and Wisconsin in 2004 and 2000 that were tested for LACV infection*. ..... 51
Table 2.3: Nonsynonymous mutations found in sequences of LACV RNA amplified from field collected mosquitoes* ..... 54
Table 2.4: LACV segment reassortment occurs in field collected mosquitoes as revealed by a linkage disequilibrium analysis* ..... 62
Table 2.5: LACV segment reassortment occurs in field collected mosquitoes as revealed by a linkage disequilibrium analysis using the conservative analyses* ..... 63
Table 2.6: The S segment is the segment likely to reassort revealed by the linkage disequilibrium analysis of haplotype combinations of S*M, M*L, and S*L* ..... 64
Table 2.7: Minimal linkage disequilibrium was observed in the 2x2 linkage disequilibrium analysis of haplotype combinations of S*M, M*L, and S*L using the conservative clades.* ..... 64

Table 3.1. LACV-infected field collected mosquitoes from Wisconsin, Minnesota, and Iowa collected in 2000, 2004 and 2005* ..... 83
Table 3.2: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all nine counties ..... 87
Table 3.3: Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all nine counties ..... 87
Table 3.4: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all four quadrants ..... 88
Table 3.5: Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all four quadrants ..... 88
Table 3.6: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all four environmental terrain categories ..... 90
Table 3.7 Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all four environmental terrain categories ..... 90
Table 3.8: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all three years ..... 91
Table 3.9: Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all three years ..... 91
Table 3.10: Evidence for quasispecies in LACV S and M segments in all 3 years LACV RNA was analyzed. ..... 96

Table 4.1: Primer sets for amplification of the three segments of LACV ..... 116
Table 4.2: Virus RNA sequences used in maximum parsimony analysis of virus isolates ..... 119

Table 4.3: MAS used in maximum parsimony analysis* ..... 120
Table 4.4: Prevalence and distribution of LACV normally-infected and LACV super-infected mosquitoes in the 2006 and 2007 collections ..... 123
Table 4.5: Prevalence of super-infected mosquitoes by collection site in 2006 and 2007 ..... 123
Table 4.6: LACV titer of super field-infected mosquitoes ..... 124
Table 4.7: Nucleotide diversity of the S segment comparing super-infected isolates and previously published sequences from the similar collection region ..... 137
Table 4.8: Nucleotide diversity of the M segment comparing super-infected isolates and previously published sequences from the similar collection region ..... 137
Table 4.9: Nucleotide diversity of the L segment comparing super-infected isolates and previously published sequences from the similar collection region ..... 137

Table 5.1: Mosquitoes and conditions used for insemination and cuticular hydrocarbon experiments ..... 154
Table 5.2: There is no association between genome equivalents and insemination ..... 162
Table 5.3: CHCs of adult female *Aedes triseriatus* ..... 164

## TABLE OF APPENDICES

Appendix 1: S segment sequence alignment of LACV prototype and virus isolated from super-infected mosquitoes.....179
Appendix 2: M segment sequence alignment of LACV prototype and virus isolated from super-infected mosquitoes.....182
Appendix 3: L segment sequence alignment of LACV prototype and virus isolated from super-infected mosquitoes.....191

# I. LITERATURE REVIEW

## A. INTRODUCTION

Diseases caused by arthropod-borne viruses (arboviruses) are resurgent worldwide and the reasons for this are multi-factorial. Examples of these factors and resurgent diseases will be reviewed. La Crosse virus is a prototypical emerged arbovirus. In this dissertation, factors that condition the evolutionary, epidemic and maintenance potential of this arbovirus are investigated and provide insight into the determinants of arbovirus emergence and epidemic potential.

### 1. The global resurgence of arbovirus diseases

The resurgence of arboviruses and resultant diseases is a significant public health threat, as demonstrated by the dramatic increase in emerging or remerging diseases worldwide. The majority of the newly recognized viral diseases of humans are zoonoses that have other animal reservoirs and which, due to focal and/or global environmental, societal, and demographic changes, were able to jump species to infect humans (Gubler, 2002). The epidemic resurgence of many well-known arboviruses that were thought to be under control or insignificant, such as dengue virus, West Nile virus (WNV), Japanese encephalitis, yellow fever, Rift Valley fever, Venezuelan equine encephalitis and Ross River, has affected both animals and humans worldwide (Gubler, 2002). In the United States the emergence of WNV and the apparent expansion of La Crosse virus (LACV) into the southern United States and transmission by a new mosquito vector, *Aedes albopictus,* are of public health concern.

In the past twenty years, epidemic arboviral activity has increased globally due to a variety of reasons, including urbanization, increased population, transportation, lack of mosquito control, increased geographic distribution of vectors, and evolution of

arboviruses through genetic drift and genetic shift (Gubler, 2002). Viruses that are introduced into new geographic areas may be able to take advantage of susceptible vertebrate and arthropod hosts, resulting in major epidemics/epizootics, such as that observed following the introduction of WNV into the United States (Gubler, 2002). Many arboviruses have also expanded their geographic distribution regionally, which can also result in major epidemics, e.g. Rift Valley fever, Ross River, Japanese encephalitis, and Venezuelan equine encephalitis (Gubler, 2002). Increased distribution of some mosquito vectors, for example *Aedes albopictus*, may result in a global expansion of viruses, such as dengue virus (Gubler, 2002, Tatem et al., 2006). Humans have increasingly inhabited wooded areas thereby impinging upon the vector-host transmission cycles and resulting in viral infections such as with LACV (Rust et al., 1999). Humans are typically dead end hosts for these viruses, but if the viruses adapt to human hosts there is the potential for the emergence of new cycles, causing epidemics. Viruses with segmented genomes, such as those in the *Bunyaviridae* family, have the potential of evolving through segment reassortment and intramolecular changes and thus possess significant epidemic potential.

The decline in public health infrastructure in many countries in the past twenty years is also contributing to the resurgence of arboviruses. There is a lack of funding for the development of effective prevention and control programs for vector-borne diseases. There is also a need to change public health policy to place more emphasis on preventing epidemics rather than on emergency response (Gubler, 1998).

## 2. Emergence and resurgence of selected arboviruses

### a. La Crosse virus – La Crosse encephalitis

In the 1970s, LACV (Family: *Bunyaviridae*, Genus: *Orthobunyavirus*, California serogroup) emerged as a significant human pathogen in the upper Midwest of the United States possibly due to increased interaction with humans in forested areas (Beaty et al., 2000). LACV was first isolated from the brain of a four year-old girl who acquired an infection in Minnesota and died in La Crosse, WI, in 1960 (Thompson et al., 1965). LACV is the most significant virus in the California serogroup because it is the primary cause of pediatric encephalitis and aseptic meningitis in the United States (Beaty et al., 2000, Rust et al., 1999). The incidence rate of LACV infections is five to ten cases per 100,000 persons in endemic areas, with typically 80 cases reported per year (Beaty et al., 2000, Borucki et al., 2002, Tsai, 1991). Estimates of the ratio of mild or subclinical LACV infection to encephalitic LACV cases range from 2:1 to 1500:1, depending upon age, and therefore, as many as 300,000 mild LACV infections might occur each year in the United States (Grimstad et al., 1984, Kappus et al., 1982, Monath et al., 1970). Most LACV infections are probably ascribed to "flu" or "summer colds". Neuroinvasive LACV disease is primarily observed in children, where fatalities are rare and persistent cognitive impairment and seizures are possible sequelae (McJunkin et al., 2001).

LACV infection has become an emerging disease in the southeastern portion of the United States (North Carolina, Tennessee, West Virginia, Kentucky) in the last ten years (Gerhardt et al., 2001, Jones et al., 1999). From 1963 to 1996, nine cases of LACV encephalitis were reported in TN and three in NC. From 1997 to 1999, 26 pediatric LACV encephalitis cases were reported in eastern TN and southeastern KY (Jones et al.,

1999) and 25 cases in NC (Kappus et al., 1982). The recent introduction of *Aedes albopictus* is thought to be a possible reason for the increase in LACV cases in the southeastern United States. LACV was isolated from naturally infected *Ae. albopictus* mosquitoes and epidemiological studies have demonstrated that the presence of *Ae. albopicutus* mosquitoes is a major risk factor for infection. *Ae. albopictus* is an aggressive daytime feeding mosquito that feeds on a number of hosts including humans, domestic and wild animals, and birds (Gratz, 2004). These feeding habits could potentially increase the number of human cases in endemic foci or expand the range of the disease.

**b. Dengue virus – Epidemic dengue and dengue hemorrhagic fever**

The incidence of dengue infections has increased significantly over the past decades. Dengue has been recognized in over 100 countries, placing nearly 2.5 billion people at risk for infection. Every year there are an estimated 50-100 million cases of dengue fever and about 250,000 – 500,000 cases of dengue hemorrhagic fever (DHF) worldwide (Calisher, 2005, WHO, 1999). Most cases of DHF and dengue shock syndrome (DSS) are of sufficient severity to require hospitalization, and case fatality rates can be as high as 5% (WHO, 2002).

Multiple factors are thought to contribute to the recent emergence of epidemic dengue and DHF/DSS in the Americas, including both demographic and societal changes. The increasing human population, especially in developing countries, is resulting in more susceptible individuals, overpopulation, and overcrowding in the cities, leading to poor environmental sanitation and living conditions (Gubler, 2004). Poor housing, sewer and waste management systems, and lack of a reliable water source make

it necessary to collect and store water, thus creating ideal places and prime conditions for breeding of the anthropophilic *Ae. aegypti* mosquito (vector of dengue virus) (Gubler, 2004). The geographic distribution of *Ae. aegypti* mosquitoes has increased due to urbanization, which has also resulted in more breeding sites (Gubler, 2004). Increased air travel enables the introduction, transport and exchange of dengue viruses into and between population centers of the tropics (Gubler, 1998).

**c. West Nile virus – West Nile fever and encephalitis**

The mid-1990s saw a change in the epidemiology of West Nile virus (WNV) (family: *Flaviviridae*, genus: *flavivirus*). Epizootics and epidemics of severe neurologic disease in horses, birds and humans began to occur in many parts of the world (Hubalek & Halouzka, 1999, Marfin & Gubler, 2001). Apparently a new virulent WNV strain with greater epidemic potential emerged in Europe, Africa, and the Middle East and spread to new geographic areas including the United States in the late 1990s (Marfin & Gubler, 2001). The WNV strain introduced into the United States in 1999 is thought to have originated in the Middle East. Genetic sequence studies revealed that the United States WNV strain is most similar to the virus strain that caused an epizootic in Israel in 1998 (Lanciotti et al., 1999). There are multiple theories about how WNV was introduced into the United States. One theory is that illegal trafficking of exotic birds may have introduced the virus. Other possibilities are that an infected mosquito may have been on an airplane from the Middle East, the virus could have been introduced on purpose, or an infected migrating bird from Africa may have become disoriented and ended up in New York City (Gubler, 2002). The introduction of WNV into the Western Hemisphere has had a great public health and economic impact in the United States in terms of human

infections. Between 1999 and 2005, there were 18,700 human cases of WNV infection in the United States (Sejvar, 2006) as the virus swept through a susceptible population. Fever and other syndromes occurred in 10,724 cases and neuroinvasive disease was observed in 7,874 cases (Sejvar, 2006). The introduction and spread of WNV to the Western Hemisphere is an example of how important emerging and reemerging vector-borne diseases have become in today's society.

**d. Other arboviruses of public health emergence concern in the United States**

Other arboviruses such as Eastern Equine encephalitis virus (EEE, Family: *Togaviridae,* Genus: *Alphavirus*), Western Equine encephalitis virus (WEE, Family: *Togaviridae,* Genus: *Alphavirus*) and St. Louis encephalitis virus (SLE, Family: *Flaviridae*, Genus: *Flavivirus*) continue to be of public health concern in the United States. EEE, WEE, and SLE have caused significant human illness (Karabatsos, 1985. 2001 update). WEE and EEE have caused epizootics since the 1930s (Meyer et al., 1931, Ten Broeck & Merrill, 1933). SLE was recognized as a cause of human encephalitis during an outbreak in Paris, IL in 1932 (Monath, 1990). Since 1964, SLE has been responsible for at least 4,600 cases of severe illnesses and many mild or subclinical infections (Chamberlain, 1980, Reisen, 2003).

**B. ARTHROPOD-BORNE VIRUSES**

**1. Arthropod-borne (Arbo) virus transmission cycles**

Arboviruses are naturally maintained in cycles by hematophagous arthropod vectors that biologically transmit virus between vertebrate hosts. Vectors can become infected by a variety of mechanisms, depending on the system, including ingestion of viremic blood from a vertebrate host, transovarial transmission, venereal transmission

during mating and even vertical transmission of the virus to embryos during oviposition (Higgs & Beaty, 2005). Arboviruses must be capable of infection and replication in two phylogenetically different systems, the invertebrate vector and the vertebrate host.

**2. Vector competence**

Vector competence is the intrinsic ability of a vector to biologically transmit a disease agent (Higgs & Beaty, 2005). This includes susceptibility to infection, permissiveness for pathogen reproduction/development, duration of the extrinsic incubation period (EIP: the period from ingestion of the infectious blood meal to transmission capability), and transmission efficiency (Mitchell, 1983). A competent vector species has a low infection threshold, is permissive for viral multiplication, and should sustain little or no reduction in fitness from viral infection. The shorter the EIP, the more competent the vector will be because of the potential to encounter more hosts after becoming infected.

Vector competence could be reduced by multiple barriers to infection in the midgut, salivary glands and ovaries (Hardy, 1988, Mitchell, 1983). The first barrier a virus may encounter is the midgut infection barrier. A midgut escape barrier occurs when the midgut is infected but the virus cannot successfully disseminate to infect other organs.

There are a variety of hypotheses about the existence of barriers to midgut infection (Higgs & Beaty, 2005). It has been suggested that the formation of the peritrophic matrix may prevent the virus from contacting and infecting the midgut cells (Devenport & Jacobs-Lorena, 2005). It is also possible that refractory arthropods possess specific defensive mechanisms to protect against viral infections. Mucopolysaccharides or other antiviral agents could be secreted in the midgut. RNA interference (RNAi) also

gives vectors the ability to defend against a viral infection (Sanchez-Vargas et al., 2004). Viral inactivation or inappropriate processing of viral surface glycoproteins by midgut enzymes is another potential explanation for the midgut infection barrier. An additional hypothesis is that refractory arthropods do not have the specific receptors on midgut epithelial cells for viral attachment (Higgs & Beaty, 2005).

Salivary gland infection and escape barriers and ovarian infection barriers also may occur. The salivary glands either do not become infected or can become infected but the virus is not transmitted to a vertebrate host. Little is known about the molecular bases of ovarian infection barriers. When developing follicles do become infected, the virus must replicate and persist through embryogenesis, larval development, and metamorphosis for the adults of the successive generation to be infected (Higgs & Beaty, 2005).

**3. Vectorial capacity**

Vectorial capacity is the overall ability of a vector species in a given location at a specific time to transmit a pathogen (MacDonald, 1957). Quantitatively, vectorial capacity can be defined as the number of infectious bites a person (or vertebrate host) receives daily. The vector population size, vector longevity, length and number of gonadotrophic cycles, feeding behavior, and diel activity are all factors that affect the vectorial capacity in a population. The vectorial capacity includes the vector interactions with the pathogen and the vertebrate host and vector characteristics not directly related to either the pathogen or vertebrate host.

## 4. Mechanical, horizontal and vertical transmission of arboviruses

Vectors may transmit infectious agents biologically or mechanically (Chamberlain & Sudia, 1961). For biological transmission, the pathogens reproduce or develop in the arthropod vector before being transmitted to the next vertebrate host. In mechanical transmission, the pathogen does not reproduce or develop in the vector; the pathogen is simply physically transmitted from one vertebrate host to another (Chamberlain & Sudia, 1961). Typically the mosquito's mouthparts become contaminated with pathogens that are then inoculated into another vertebrate host during the next feeding attempt.

Arboviruses are biologically transmitted by vectors through both horizontal transmission (transmission between vectors and vertebrate hosts or venereal transmission between vectors during mating) and vertical transmission (transmission from an arthropod to the progeny) (Chamberlain & Sudia, 1961, Higgs & Beaty, 2005).

Arboviruses must infect and replicate in the vector to be horizontally transmitted to a new vertebrate host. During the EIP, the virus infects and replicates in the midgut epithelial cells and then disseminates to secondary target organs, possibly through the hemocoel. Virions then disperse in circulating hemolymph. Once the salivary glands become infected and shed virus into the salivary ducts, the virus can be transmitted to vertebrates during blood feeding. The length of the EIP depends on the temperature, specific vectors and viruses. A typical EIP is 10-14 days for a bunyavirus and 6-7 days for an alphavirus (Higgs, 2004). Once a vector is infected with a virus, it remains infected throughout its lifespan, and the vector can potentially infect hosts each time it feeds. For the transmission cycle to remain intact, arboviruses must cause a high-titered viremia in

the natural vertebrate host. High-titered viremias of long duration maximize the opportunity for vector infection and for transmission of the virus.

Arboviruses are vertically transmitted when the arthropod vector passes the pathogen to its progeny. Most often, a female arthropod infects her progeny through a transovarial or transovum route. California serogroup viruses of the *Bunyaviridae* family are efficiently transmitted transovarially and the follicles are permissive to virus replication (Watts et al., 1973). Flaviviruses can be vertically transmitted during oviposition when the virus is transferred into the egg through the micropyle as the egg passes through the oviduct (Aitken et al., 1979). Males can vertically infect progeny through the seminal fluids and can venereally transmit arboviruses during copulation.

**5. Virus adaptation in an arthropod host**

Arboviruses acquire many adaptations that are beneficial to their own survival and dispersal (Black & Salman, 2005). These include adaptations that promote optimum transmission of the arbovirus. Once the arbovirus gains access to the arthropod vector, it is faced with the problem of invading the arthropod gut, disseminating to various organs, and eventually ending up in an organ of transmission without affecting the overall host fitness. Most arboviruses target the host salivary glands, but another option for transmission is the male or female reproductive system. The vector-arbovirus relationship is commonly referred to as cospeciation, which implies that arbovirus phylogenies parallel host phylogenies because they have coevolved (Brooks & Leman, 1993).

Most arboviruses evolved to infect, replicate, and be transmitted without being detected or by having a very small impact on the host survival. A competent vector is permissive to an arbovirus for replication, survival and transmission and does not mount

an effective immune response against the arbovirus. Evasion of the innate immune system of the arthropod is imperative. Plasticity in vector competence is consistent with a general genetic model in which multiple factors must be present in the arthropod for successful transmission of an arbovirus.

Part of the definition of an arbovirus is no causation of untoward effects on the natural vector. However, mortality has been shown to be higher in mosquito eggs transovarially infected with LACV than in uninfected eggs when the eggs overwinter and emerge from diapause (McGaw et al., 1998). Recent studies have also demonstrated cytopathology of alphaviruses and WNV on the natural mosquito vector (Girard et al., 2005, Girard et al., 2007, Scott & Lorenz, 1998, Scott & Weaver, 1989, Weaver et al., 1988).

It has also been documented that arboviruses affect arthropod behavior. There is an affect on both longevity and spontaneous flight of *Culex tarsalis* infected with WEE virus (Lee et al., 2000). Mosquitoes infected with WEE had flight activity scores that were 27.5% lower than uninfected mosquitoes. Another study documented the effects of dengue virus infection on feeding behavior (Platt et al., 1997). The mean total time required for feeding by infected mosquitoes was significantly longer than the time required by uninfected mosquitoes. Similarly, the mean time spent probing was significantly longer for infected mosquitoes than for uninfected mosquitoes, which could promote transmission.

LACV infection causes changes in behavior in *Ae. triseriatus* mosquitoes (Grimstad et al., 1980). Infected females tend to probe more and engorge less than uninfected females. For example, 21% of infected females took a partial bloodmeal with

one probe compared to 52% of uninfected females that were fully engorged with one probe. Seventy-nine percent of the infected females made multiple probes to obtain a partial bloodmeal compared to 48% of the uninfected females making multiple probes to feed. This study provides evidence that LACV may have a deleterious effect in *Ae. triseriatus* mosquitoes but enhances virus transmission.

## C. *BUNYAVIRIDAE*

### 1. *Bunyaviridae* family overview

La Crosse virus (LACV) is a member of the *Bunyaviridae* family, which is the largest family of arboviruses with >250 named viruses (Beaty & Calisher, 1991). The family is comprised of a diverse group of RNA viruses made up of five genera, *Orthobunyavirus, Hantavirus, Phlebovirus, Nairovirus,* and *Tospovirus*. The genera are distinguished by antigenic, serological, molecular, and structural differences (Calisher, 1994). Members of the *Bunyaviridae* family are found worldwide and are significant pathogens of humans, domestic and wild animals (*Orthobunyavirus, Hantavirus, Phlebovirus, and Nairovirus)* and plants (*Tospovirus)*. Most of these viruses are transmitted between an arthropod vector (mosquitoes, sand flies, and ticks) and a vertebrate host. Hantaviruses, however, do not infect arthropods and are maintained in nature through a rodent-rodent cycle with persistent infection of a rodent host. Non-vectored transmission has also been observed for some *Nairoviruses* (Crimean Congo hemorrhagic fever) and *Phleboviruses* (Rift Valley fever). These viruses can be spread to humans by exposure to infected animal tissues or body fluids (Soldan & Gonzalez-Scarano, 2005).

## 2. *Orthobunyavirus* genus

*Orthobunyavirus*, the largest genus of the *Bunyaviridae* family, is comprised of 172 individual viruses classified within 18 serogroups based on serological and molecular relationships (Soldan & Gonzalez-Scarano, 2005). Most members of the genus are transmitted by mosquitoes and many have been associated with human and animal infections. The California serogroup has been the most studied because of some important members such as LACV (Soldan & Gonzalez-Scarano, 2005). Other viruses in this family include California encephalitis, Jamestown Canyon, Snowshoe hare, and Inkoo. California serogroup viruses are found in a wide range of climates from tropic (Melao and Serra do Navio) to temperate (LACV) to artic (Inkoo and Snowshoe Hare virus) (Beaty & Calisher, 1991, LeDuc, 1979).

## 3. Molecular biology and gene structure-function relationships of *Bunyaviruses*

The *Bunyavirus* genome is negative sense and tripartite. The three RNA segments designated large (L), medium (M) and small (S) encode a variety of proteins. The coding strategy differs somewhat between the genera within the family (Elliott, 1990). The L segment, ranging in size from 6.4 kb (*Phleboviruses*) to 12.2 kb (*Nairoviruses*), encodes the RNA-dependent RNA polymerase (Endres et al., 1989). The M segment, ranging from 3.6 kb in *Hantaviruses* to 4.9 kb in *Nairoviruses*, encodes a precursor polypeptide that is post-translationally cleaved to generate the G1 and G2 glycoproteins (Elliott, 1985, Fuller & Bishop, 1982, Gentsch & Bishop, 1979). The S segment, varying from ~1 kb (*Orthobunyaviruses*) to 3 kb (*Tospoviruses*), encodes the nucleocapsid protein (Cabradilla et al., 1983, Fuller & Bishop, 1982). Nonstructural proteins are encoded in the M segment of *Phleboviruses* and *Orthobunyaviruses* and in the S segment of

*Orthobunyaviruses* (Elliott, 1985, Fuller & Bishop, 1982). The termini of the 3' and 5' non-coding regions of each segment are complementary and highly conserved, forming a helical nucleocapsid (Obijeski et al., 1980).

Neuroinvasiveness and mosquito infectivity of California serogroup viruses are found to be associated with the M segment (Gonzalez-Scarano et al., 1988, Gonzalez-Scarano et al., 1992, Janssen et al., 1986). Genome structure-function studies using reassortant viruses containing segments from both a virulent California serogroup virus (LACV) and an avirulent virus (Tahyna) revealed the viral genetic determinants of neuroinvasiveness. Viruses containing the M segment from LACV were significantly more neuroinvasive than viruses containing the M segment from the Tahyna virus. Comparison of four different biological markers (neuroinvasiveness and neurovirulence in mice; oral and intrathoracic infectivity in mosquitoes) were used to determine which segment is most important in virulence and infectivity (Beaty et al., 1983, Beaty et al., 1981, Beaty et al., 1982, Bishop, 1979, Bishop et al., 1980, Gonzalez-Scarano, 1985, Gonzalez-Scarano et al., 1987, Shope et al., 1981). These gene structure-function studies revealed that the M segment is a major determinant of mouse virulence and mosquito infectivity and that the G1 glycoprotein influences virulence tropisms and may be its major determinant. The four biological markers can vary independently, which suggests that there may be multiple sites within the M segment RNA or contributions from other segments that can influence virulence (Gonzalez-Scarano et al., 1988). The S and L segments may, on certain genetic backgrounds, modulate the dominant effect of the M segment (Janssen et al., 1986). Reassortants with avirulent Tahyna virus M RNA segments and virulent LACV S and L segments were about ten-fold more virulent than

the parent Tahyna virus. These results reveal that virulence is usually under polygenic control (Janssen et al., 1986).

LACV infects both vertebrate and mosquito cells through an interaction between the G1 and G2 glycoproteins and the cell surface. G1 covers G2 on the virion surface. G1 is the ligand for cellular receptors on vertebrate cells and is the immunodominant target for most neutralizing monoclonal antibodies (Gonzalez-Scarano et al., 1982). G2 is the ligand for an unknown receptor on the cell of the mosquito midgut (Ludwig et al., 1991). Ingested virus can bind to the midgut after mosquito proteases cleave the G1 protein, thus exposing G2 and allowing virus to attach to specific receptor proteins or fuse to cells (Ludwig et al., 1989, Ludwig et al., 1991). The virus may then enter the cells through endocytosis or by fusion with the plasma membrane releasing the genome into the cell. Acidification in the lysosome follows resulting in a conformational change in G1 and/or G2 that facilitates fusion of the viral and cellular membranes (Schmaljohn & Hooper, 2001). Once the viral genome is uncoated, primary transcription of the negative sense genome begins by the virion-associated polymerase using the three RNA templates, where only ribonucleocapsid can serve as transcription templates (Schmaljohn & Hooper, 2001). The viral polymerase contains a methylated cap-dependent endonuclease activity that presumably cleaves 10-18 capped nucleotides from the 5' termini of cellular mRNA. These caps are used as primers for the viral mRNA (Kolakofsky, 1991). There is a switch from making truncated positive sense mRNA to making full length positive RNA for replication following an accumulation of the N protein for binding of newly synthesized genome segments. The mRNAs of the S and L segments are processed on free ribosomes, whereas the mRNAs of the M segment are processed on membrane-bound ribosomes

(Schmaljohn & Hooper, 2001). Virions are formed intracellularly by a budding process at smooth surface vesicles in the Golgi and then transported to the cell surface within vesicles analogous to the secretory pathway (Schmaljohn & Hooper, 2001).

## D. LA CROSSE VIRUS

### 1. La Crosse virus epidemiology and clinical manifestations

LACV is found primarily in the upper Midwestern United States notably in Wisconsin, Minnesota, Ohio, Illinois, Indiana and Iowa. Most cases of LACV infection occur between July and September (94%), when *Aedes triseriatus* mosquitoes are most active (Rust et al., 1999). More than 95% of LACV cases in Wisconsin, Ohio, and Minnesota occur in children (between 0.1 and 17 years) (Gunderson & Brown, 1983). Children aged between 5 and 10 account for 45%-62% of the cases, adults account for 3% and children aged 0.1 to 1 year account for less than 2% of cases (Gunderson & Brown, 1983, Kappus et al., 1983). The age-related factors that condition disease severity remain unknown.

There are multiple risk factors for LACV infection. Most children who develop LACV encephalitis live on farms or in forested areas (Rust et al., 1999). Proximity to oviposition sites for *Aedes triseriatus* appears to be an important risk factor as well as the number of hours spent outdoors during daylight (Erwin et al., 2002, Woodruff et al., 1992). Boys are more likely to become infected with LACV due to increased time spent outside playing near woods, climbing trees, or building tree houses (Chun et al., 1968, Gunderson & Brown, 1983).

Children with LACV encephalitis or aseptic meningitis typically have malaise, restlessness, and irritability for one to three days before the onset of fever. Fever and

headache occur in most cases along with sensory disturbance, ranging from confusion to coma. Single seizures occur in about 42%-60% of the cases and recurring seizures occur in 10-15% of cases (case fatality rate in these patients is one percent)(Balfour et al., 1973, Beaty et al., 2000, Gunderson & Brown, 1983). There are lasting sequelae for some cases. Epilepsy is the most serious consequence of LACV infection, developing in 13-28% of cases (Rust et al., 1999). Migraine headaches, learning difficulties, and behavioral disturbances are also known to result (Chun et al., 1968). Children who recover from severe LACV encephalitis may have a significantly lower IQ than average and a high prevalence (60% of those tested) of attention-deficit-hyperactivity disorder (McJunkin et al., 2001).

**2. La Crosse virus transmission cycle**

LACV is maintained in nature in a cycle between *Aedes* (*Ochlerotatus*) *triseriatus* mosquitoes and chipmunks or tree squirrels (Rust et al., 1999). *Ae. triseriatus* is a treehole mosquito that thrives in forested areas with deciduous oak and hickory hardwoods, a habitat frequently found in the upper Midwest, Ohio, southern Indiana, and West Virginia. The combination of both the *Ae. triseriatus* mosquito and the vertebrate hosts, chipmunks and tree squirrels, in the forested habitat, determines the viral distribution (Rust et al., 1999).

Summer amplification of LACV is a significant process in maintaining the transmission cycle. Several common species of forest mammals, particularly the eastern grey squirrel (*Sciurus carolenensis*), western fox squirrel (*Sciurus niger*) and chipmunk (*Tamias striatus*) are considered important vertebrate hosts. Anti-LACV antibody prevalence rates have been reported as high as 100% in adult and spring-born juveniles of

these species by September (Gauld et al., 1974). Serological evidence for natural LACV infection has also been demonstrated in 58% of free ranging foxes (*Vulpes fulva*), 57% of gray foxes (*Urocyon cinereoargenteus*) and 25% of raccoons (*Procyon lotor*) (Amundson & Yuill, 1981). Viremia titers for chipmunks vary between 3.9-5.7 $\log_{10}$SMICLD$_{50}$/1ml with the average duration of viremia about 2.4 days (Pantuwatana et al., 1972, Patrican et al., 1985). Chipmunks develop a high enough viremia to transmit LACV to a susceptible mosquito. Studies have indicated a continual recruitment of susceptible chipmunks throughout the summer. Spring juveniles are born between mid-April and mid-May and the summer juveniles are born between mid-July and late August. Summer juveniles do not come above ground until September and have a protective passive maternal antibody when they do come above ground, and become susceptible adults the following year. This ensures continuous transmission of LACV. Clearly small mammals are important reservoir hosts that contribute to the transmission cycle of LACV in nature.

**3. La Crosse virus pathogenesis in vertebrate hosts**

The pathogenesis of LACV infection in either a chipmunk/tree squirrel or a human child begins similarly. The host becomes infected through the bite of an infected mosquito (Watts et al., 1972). The mosquito injects the virus from its salivary glands into the host's skin. The virus infects the skeletal muscle where it begins its initial replication. Virus replication increases the titer of the virus and after a period of time the virus enters the blood, either directly or through the lymph system. This enables the virus to gain access to a variety of tissues including muscle, connective tissue, liver and spleen to continue replication. Viral replication increases the titer of the virus and after a period of time the virus enters the blood causing the secondary viremia. The titer of the virus in the

blood during the secondary viremia is sufficient to infect biting mosquitoes (Janssen et al., 1984, Johnson, 1983, Parsonson & McPhee, 1985). In small mammals, there is minimal fitness cost to the reservoir host and the immune response enables clearing of the virus from the reservoir host with minimal pathogenesis (Higgs & Beaty, 2005, Seymour et al., 1983).

In a susceptible child, the secondary viremia allows the virus to infect the target organ, the brain. LACV causes encephalitis in children after crossing the blood brain barrier, and infecting the neurons and causing necrosis and apoptosis. The presence of the virus in the brain triggers an inflammatory response that results in infiltration of lymphocytes into the brain (Kalfayan, 1983). The human is a dead end host because viremia titers are not sufficient to infect susceptible mosquitoes.

**4. *Aedes triseriatus*-LACV interactions**

*Aedes triseriatus* mosquitoes become infected following ingesting of an infectious bloodmeal from a viremic vertebrate host. In the mosquito, the virus infects epithelial cell layer of the midgut (Beaty et al., 1981, Sundin et al., 1987). Replication results in the dissemination of the virus across the basal lamina and into the hemocoel, which allows the virus to travel to and replicate in other tissues, such as the heart, neural ganglia, fat body, ovaries and salivary glands. The salivary glands are the last organ infected at 7-14 days post-infection (Beaty & Calisher, 1991, Schmaljohn, 1996). Once the virus is in the salivary glands, it replicates to high titer and is transmitted to the vertebrate host in the mosquito saliva. LACV analyte was detected in the ovaries two days after infection suggesting that LACV can infect the ovaries before dissemination of the virus from the midgut (Chandler et al., 1998). Infection of the ovaries results in transmission of LACV

to the progeny. LACV accumulates in all tissues of the ovaries and maintains a persistent infection that allows efficient transovarial transmission (TOT) of the virus. The eggs in the first ovarian cycle are not infected (Miller et al., 1979). By the time the virus infects the follicular epithelium of the ovaries, it is too late to infect the first cycle of eggs, although a few eggs may be infected by early virus infection of the tissues such as calyx (Chandler et al., 1998). Eggs in subsequent ovarian cycles are efficiently infected with LACV.

*Aedes triseriatus* overwinters solely as diapausing eggs in the northern part of the United States (Pantuwatana et al., 1974, Watts et al., 1974). Diapause is induced when the fully formed embryo within the eggshell is exposed to a short day length (Shroyer & Craig, 1980). Diapause could protect infected embryos when ambient temperatures would otherwise permit host metabolic activity and virus replication (Borucki et al., 2002).

Watts *et al.* were the first to find that mosquitoes were infected with LACV without a prior infectious bloodmeal (Watts et al., 1974). This was the first conclusive evidence for the role of TOT of an arbovirus in an overwintering mosquito. TOT is an extraordinary method of amplification of the virus since each female mosquito produces many eggs. Several years of dormancy can elapse before conditions are favorable for hatching diapaused eggs, which assures maintenance of the virus in nature over many years (Rust et al., 1999).

Newly hatched females are then capable of transmitting the virus upon emergence. The TOT rate (percentage of infected females that transmit virus to their progeny) and filial infection rate (the percentage of infected larvae from a single LACV-infected female) in laboratory conditions can each exceed 70% (Beaty & Bishop, 1988).

Different geographic strains of *Ae. triseriatus* vary in their ability to vertically transmit LACV (Miller et al., 1982). Eggs from different states (Wisconsin, Florida and Connecticut) have significantly different filial infection rates - 50%, 30% and 20%, respectively - which could have a significant epidemiological relevance in terms of LACV endemicity (Miller et al., 1982, Woodring et al., 1998).

There is evidence for a negative fitness effect of TOT on the overwintering success of *Ae. triseriatus* mosquitoes. Embryos from LACV TOT+ and TOT- *Ae. triseriatus* colonies were induced into diapause and shipped to Wisconsin in 1993 for overwintering (McGaw et al., 1998). A comparison of mortality rates, diapause status, and infection rates revealed that LACV infection does affect survival. A greater proportion of uninfected eggs successfully overwintered than the infected eggs. The mortality in LACV-infected eggs occurred after emergence of embryos from diapause in the spring (McGaw et al., 1998). There was no difference between infected and uninfected mosquitoes in survivability of embryos in diapause. This suggests that the diapause condition attenuated deleterious virus effects on embryos and that diapause intensity and duration could condition the efficiency of both vector and virus overwintering.

LACV can also be transmitted venereally between an infected male and a susceptible female during copulation (Beaty & Calisher, 1991, Thompson & Beaty, 1977, Thompson & Beaty, 1978).The virus has been observed in the bursa copulatrix of 53% of female *Ae. triseriatus* mosquitoes within 24 hours after mating, but in only 3.4% of females 3-21 days post-mating (Thompson, 1979, Thompson & Beaty, 1978). Disseminated infection of female *Ae. triseriatus* mosquitoes occurred more frequently

when they were inseminated by infected males after the females had a bloodmeal (49%) than when mating took place before engorgement (4%)(Thompson, 1979). The mechanism for enhanced dissemination is unknown. Venereal transmission is not nearly as efficient as TOT, but since TOT is not 100% efficient, other forms of transmission are necessary to maintain the virus.

## E. QUASISPECIES – A MODEL FOR ARBOVIRUS TRANSMISSION CYCLES

The quasispecies model is an excellent way to explain the ability of arboviruses to pass rapidly and continuously between vectors and vertebrate hosts. Passage of arboviruses in laboratory systems rapidly selects for specific virus subpopulations that are more fit in the respective systems (Barrett et al., 1990, de la Torre et al., 1988, Novella et al., 1995). The quasispecies model hypothesizes that virus populations exist as dynamic distributions of nonidentical but related replicons (Bardos, 1965, Danielova, 1974, Eigen, 1996, Feuer et al., 1999, Mumford, 2007). A large number of different viral particles produced during an infection of a host represent a balance between the expansive influence of mutation and the conservative forces of selection (Holland et al., 1982). For a given RNA virus population, the genome sequences are similar, but virtually each genome can be unique. If random mutations have a selective advantage in terms of viral fitness (ability to replicate within the host and transmit and spread in a population) or avoidance of immune response (ability to avoid neutralization by antibody generated by earlier related strains) then these mutations may be fixed in the population (Boni et al., 2006). For example, viruses with greater infectivity have a selective advantage as they are more frequently transmitted. Viruses with altered antigenic sites, particularly those involved in virus-cell attachment, may be capable of avoiding neutralizing antibody

present in a vertebrate host resulting from a previous infection. A quasispecies that is the most fit in a vertebrate host may not be as fit in a vector, resulting in selection for new mutants.

There are a variety of stages in mosquito infection and transmission where a possible bottleneck or quasispecies could occur. When female mosquitoes take an infectious bloodmeal from a vertebrate host, only a small number of virions are ingested and infect midgut cells resulting in a bottleneck. Once the virus is in the midgut, a quasispecies may result from virus replication in the epithelial cells. A bottleneck will then occur during dissemination from the midgut since selection will likely allow only the most fit viruses to disseminate and infect target organs. When the virus initially infects the ovaries or salivary glands, a bottleneck will occur again. However, infection and replication in vector ovaries and salivary glands, especially long term replication in ovaries, will likely generate new quasispecies. Only the fittest viruses are then transmitted to the vertebrate host or the progeny.

## F. EVOLUTIONARY POTENTIAL OF THE *BUNYAVIRIDAE* FAMILY

### 1. Mechanism for the generation of genetic diversity

The evolutionary success of the *Bunyaviridae* family is attributed in part to its ability to adapt to different conditions through genetic drift (intramolecular genetic changes) and genetic shift (segment reassortment) (Beaty & Calisher, 1991, Beaty et al., 2000). Continuous evolution of the virus is occurring by the accumulation of point mutations and could be due to the struggle between host and virus to overcome the defense mechanism of the opponent (Weber & Elliot, 2002). RNA viruses are known to exhibit a high base substitution error frequency ($10^{-3}$) due to the poor fidelity of RNA

polymerase and the lack of proofreading enzymes (Holland et al., 1982). Key factors affecting the selection of variants relate to the virus, the vector, the host immune response, and the population size and structure (Mumford, 2007). The vector innate immune response contributes to selection of RNA variants. RNA interference (RNAi) is a vital immune pathway targeting double stranded RNA leading to the degradation of viral RNA (Hammond, 2005). RNAi reduces the consensus viral population and creates an environment where sequence diversity is beneficial. The viral variants that can evade the RNAi response will proceed to infect neighboring cells. RNA viruses also encounter the innate and adaptive immune response vertebrate host. Interferons are an important part of the innate immune response, whereas antibodies are the key to adaptive immunity. The interaction with the immune responses of both the vertebrate and invertebrate hosts leads to selective pressures and evolution.

Segment reassortment is an important method of evolution that can occur if the host is dually infected with two different viruses. The importance of segment reassortment has been observed with the constantly changing epidemiology of influenza virus. Influenza virus can acquire a gene for a completely new hemagglutinin or neuraminidase, giving rise to a novel subtype that spreads rapidly around the world as most or all humans are susceptible (Laver et al., 1984, Palese & Young, 1982).

**2. Evolutionary potential of La Crosse virus**

**a. Genetic drift**

There has been clear evidence that LACV evolves through genetic drift in nature. As evidenced by RNA oligonucleotide fingerprinting, no two isolates recovered in nature have identical genome sequences (Clerx-van-Haaster et al., 1982, El Said et al., 1979,

Klimas et al., 1981). This applies to both viruses isolated from the same place but at different times and at the same time but different places. Genotypic varieties of LACV were isolated from different geographic regions of the continental United States (Klimas et al., 1981). There were three genotypes observed through RNA oligonucleotide fingerprints; two isolated from the upper Midwest and one recovered from eastern Ohio, New York, Texas, Georgia, and North Carolina. Most of the LACV isolates have been shown to be closely related to each other, although they are also distinguishable (Bishop & Beaty, 1988). This is evidence for the continuous evolution of the virus by the accumulation of point mutations.

Arbovirus transmission cycles provide multiple opportunities for evolution of *Bunyaviruses*, through both genetic drift and genetic shift. Arboviruses can replicate in a vector throughout the vector's lifespan, providing ample opportunities for intramolecular changes in the viral genome, such as point mutations, sequence deletions, and inversions (Beaty & Calisher, 1991). In the case of TOT, the viruses may persist through generations of mosquitoes without being transmitted to a vertebrate, providing ample opportunity for evolution through genetic drift.

LACV genetic changes were monitored in the G1, G2 and N open reading frames (ORFs) after oral infection, dissemination, and TOT of the virus using three different LACV strains (Borucki et al., 2001). Genome sequence analyses of LACV in ovaries, salivary glands, and midgut determined that selection was occurring during virus passage in the vector. Most of the variation was observed in the G1 ORF and quasispecies were observed in the midgut (13 genotypes), followed by the ovaries (11 genotypes) with the least variation found in the salivary glands (9 genotypes) in a single mosquito (Borucki et

al., 2001). Genotypic and phenotypic changes in LACV are epidemiologically significant only if mutant viruses are transmitted to the vertebrate host and/or to the progeny (Borucki et al., 2001).

**b. Segment reassortment (genetic shift) of La Crosse virus**

Evidence clearly demonstrates that bunyavirus segment reassortment can occur naturally through dual infection of the vector (or reservoir host in the case of hantaviruses). New viral genotypes can be transmitted to progeny through bloodmeal and/or TOT. There is a real possibility that a new reassortant virus could be generated in nature, transmitted, and established in the original niche or a new niche in the environment (Beaty et al., 1985).

The ability of *Ae. triseriatus* mosquitoes to become dually infected increases the possibility of LACV evolution through segment reassortment. The ovaries are an important site of replication and therefore a possible site for segment reassortment. In a laboratory setting, high frequency (80%) of segment reassortment occurs between two heterogeneous but closely related bunyaviruses, LACV and Snowshoe hare virus (SSH) (Chandler et al., 1990). Reassortants were detected in both the body remnants and ovaries of dually infected mosquitoes, indicating a disseminated infection. The mosquitoes transovarially transmitted the reassorted viruses to their progeny with a filial infection rate of 14.2%. The infected progeny were then able to horizontally transmit the reassortant virus to mice.

Segment reassortment can occur in mosquitoes that have become dually infected by ingesting two different viruses simultaneously or by ingesting the second virus within two days of the first virus during interrupted feeding (Beaty et al., 1985). Twenty-five

percent of *Ae. triseriatus* dually infected with two different strains of LACV contain reassortants, and these new viruses can be transmitted by bite to vertebrates.

Segment reassortment has also been demonstrated in LACV TO-infected mosquitoes that subsequently ingest an infectious bloodmeal and become dually infected (Borucki et al., 1999). Approximately 20% of mosquitoes TO-infected with either a temperature sensitive LACV or wild type LACV became super-infected by ingesting blood meals containing wild type LACV or SSH virus. Reassortant viruses were detected in the super-infected mosquitoes (2.1%). Since TOT is a major maintenance and amplification mechanism for LACV in nature (Beaty & Thompson, 1975, Watts et al., 1974), the ability of even a small percentage of TO-infected mosquitoes to become super-infected and to generate reassortant viruses may play a role in LACV evolution (Borucki et al., 1999).

In nature, segment reassortment could occur either in the vertebrate host or in the arthropod vector (Borucki et al., 2002). However, despite high frequency of reassortment of LACV and SSH virus in vertebrate cells in vitro, reassortment between these viruses has not been detected in mice in laboratory studies (Beaty & Bishop, 1988). This is presumably due to the ephemeral nature of the infection in most vertebrate hosts due to production of antibodies and the lack of dual infection of cells (Borucki et al., 2002).

LACV reassortment has been found to occur in nature. Klimas *et al.* was the first to find evidence of segment reassortment of LACV in nature (Klimas et al., 1981). The authors identified distinct genotypic varieties of the virus isolated in different areas of the United States and showed evidence of reassortment between genotypes. The genomes of 23 isolates of LACV were analyzed by oligonucleotide fingerprinting and categorized in

terms of the degree of their RNA sequence relatedness. One genotype (denoted as type A) was isolated from mosquitoes from Wisconsin, Minnesota, Indiana, and Ohio, and a second genotype (denoted as type B) was isolated from mosquitoes from Minnesota, Wisconsin, and Illinois. A reassorted LACV isolated in Rochester, MN, contained the S segment of the B genotype, and the M and L segments of the A genotype. The genotypes A and B are sympatric, which could contribute to the opportunity for viral evolution through segment reassortment (Klimas et al., 1981).

**3. Segment reassortment of other *Bunyaviruses***

**a. Segment reassortment potential of California and Bunyamwera serogroup viruses**

The reassortment potential was investigated for other viruses in the California serogroup and Bunyamwera serogroup (Gentsch et al., 1980). SSH, LACV, Tahyna (TAH), Trivitattus (TVT), and Guaroa (GRO) viruses were studied. Reassortment occurred between the California serogroup viruses, SSH, LACV, TAH and TVT, but not with the virus from the Bunyamwera serogroup, GRO. Segment reassortment is apparently restricted to the serogroup. Although GRO is a possible distant relative of the California serogroup, it is not similar enough to result in segment reassortment (Gentsch et al., 1980).

**b. Group C viruses**

Molecular studies of the group C serogroup (*Bunyaviridae*, *Orthobunyavirus*) found the first evidence of segment reassortment in the field in the serogroup (Nunes et al., 2005, Shope, 1985). Serologic relationships of Group C viruses were determined by hemagglutination test or by complement fixation test of seven Group C viruses isolated from arthropods and vertebrates. The viruses segregated into three groups by HI tests,

which assay gene products of the M segment. However, the viruses segregated into three alternate groups by complement fixation tests, which assays principally the gene products of the S segment, suggesting reassortment of the RNA segments (Shope, 1985). Phylogenetic analysis of Group C viruses also found segment reassortment in nature (Nunes et al., 2005). Analysis of the nucleocapsid gene revealed that Group C members were distributed into three major clades. Analysis of a Caraparu viral strain (BeH 5536) revealed that it had an S segment sequence nearly identical to that of Oriboca virus, but an M segment sequence similar to Caraparu virus and is therefore a natural reassortant.

**c. Ngari virus**

Gerrard *et al.* isolated a virus from the genus *Orthobunyavirus* that was a genetic reassortant based on sequence analysis of the three genomic RNA segments (Gerrard et al., 2004). The S and L segments were derived from Bunyamwera virus but the M segment was from an unidentified virus of the genus *Orthobunyavirus*. Sequence analysis of the unknown M segment revealed that it was identical to the Ngari virus. The S and L segments showed high sequence identity with those of the Bunyamwera virus, so this isolate was a reassortment with the genotype of S, Bunyamwera; M, Ngari virus; and L, Bunyamwera virus (Gerrard et al., 2004). Further analysis revealed that the M segment from the recently identified Ngari virus has 89-95% sequence identity to Batai virus, an *Orthobunyavirus* that has not been isolated from humans (Briese et al., 2006, Yanase et al., 2006). This reassorted virus caused hemorrhagic fever outbreaks throughout sub-Saharan Africa (Briese et al., 2006, Yanase et al., 2006).

**d. Crimean Congo hemorrhagic fever virus**

Recent evidence of segment reassortment has been observed with Crimean Congo hemorrhagic fever virus (CCHF) (*Bunyaviridae*, *Nairovirus*). Analysis of the S and M segments revealed that the phylogenetic grouping of some strains differ between the two segments (Hewson et al., 2004). This analysis also suggested a possible geographical correlation between the relationships of S and L segment, but not between the M segments. Further epidemiological studies supported these results and found that closely related S and L segments must co-segregate in order to produce viable virus, restricting reassortment opportunities (Chamberlain et al., 2005). The S and L segment of CCHF have evolved together.

**e. Hantavirus**

Evidence of Hantavirus reassortment in rodents (*Peromyscus maniculatus*) was observed with two genetically similar hantavirus isolates (Li et al., 1995b). The M and L segments differed from one another by only 1%, but the S segments differed by 13%. These results are very atypical for Hantaviruses. Nucleotide sequence homologies are generally very close for the S and M segments of any two viruses of the same serotype, typically differing by no more than 2%. This data suggests that there was more than one genotype present in the *P. maniculatus* population in the collection region of California and that segment reassortment occurred.

## G. MOLECULAR EPIDEMIOLOGY: DEFINING ARBOVIRUS DETERMINANTS OF EMERGENCE AND RESURGENCE

### 1. The basis of molecular epidemiology

Molecular epidemiology and phylogenetic analysis has been used to identify different viral genotypes and to understand the evolutionary history, geographic spread and epidemic potential of many viruses. Molecular epidemiology is the use of molecular genetic and biochemical markers to detect and identify pathogen species and to genetically characterize individual pathogen isolates and strains (Black & Salman, 2005). Molecular epidemiology has been used to determine the genetic variability and transmission patterns of viruses such as SLE, dengue virus and LACV (El Said et al., 1979, Klimas et al., 1981, Rico-Hesse, 1990, Trent et al., 1981).

### 2. St. Louis encephalitis (SLE) virus

The genetic diversity of 57 SLE virus strains isolated between 1933 and 1980 from various sources and localities throughout North America was elucidated through the use of RNase T1 oligonucleotide fingerprints (Trent et al., 1981). These methods allowed the determination of the variability that exists within SLE virus isolates during epidemics and over broad geographic areas. The correlation of virulence markers and oligonucleotide fingerprint patterns, and the stability of the SLE genome with respect to time were also determined. The analysis indicated that based on similarity of oligonucleotide fingerprints, SLE isolates could be divided into genotypes representing different geographic regions in North America. The geographic varieties represented isolates from (1) Central/Atlantic states, (2) Florida, and (3) the Western United States.

### 3. Dengue virus

Molecular techniques revealed the geographic distribution and route of spread of the dengue (DEN) viruses in nature (Rico-Hesse, 1990). It was determined that a Southeast Asian genotype of dengue virus may have been imported into the Americas and the cause of severe dengue outbreaks in the early 1980s. Three DEN-2 Jamaican strains isolated in 1981 and 1982 were closely related to dengue isolated from Vietnam in 1987. This was interesting because from 1977 to 1980, Vietnam experienced severe dengue epidemics yearly and during same period the Cuban military personnel were present in Vietnam. The Cuban dengue epidemic was first detected in May 1981 which raises the possibility that a viremic human may have transported a Southeast Asian strain of dengue virus from Vietnam to Cuba. Since there was very little reported DEN-2 activity in Jamaica in 1981, the close relationship between the Jamaican strains and Vietnamese strain suggests the Asian genotype may have been imported from Cuba.

As molecular techniques advanced, more information was revealed about the dengue activity in the Americas (Rico-Hesse, 1990). Analysis of DEN-2 from four American countries (Venezuela, Brazil, Colombia, and Mexico) revealed that the imported genotype from Southeast Asia displaced the "native" American genotype, which had been associated with only a less severe disease. It is thought that the Southeast Asian genotype has supplanted the "native" American genotypes in Central and South America.

The threat of American/Asian genotype viruses has increased the molecular surveillance of dengue viruses in various Central and South American countries. Diaz *et al.* reported the circulation of two DEN-2 virus genotypes in the Yucatan: the American

genotype and the Sri Lanka genotype (Cosmopolitan genotype) (Diaz et al., 2002). DEN-2 viruses of the American-Asian genotype have also been introduced into the Yucatan (Lorono-Pino et al., 2004). DEN-2 viruses that were isolated in 2002 from the Yucatan State were most similar to the American-Asian genotype isolated in Venezuela in 1998 and Martinique in 1988. The detection of this genotype in the Yucatan State is a major concern to public health authorities since the DEN-2 virus from the American-Asian genotype is associated with more severe clinical outcomes.

A phylogenetic analysis of all four DEN serotypes in Mexico was performed with isolates from the two different epidemics in the late 1990s and early 2000s (Diaz et al., 2006). An epidemic of DHF occurred in Mexico in 1996 and 1997 and DENV-3 was the predominant serotype with other serotypes present. A second epidemic followed in 2002 and more than 1000 cases per year have been reported since. The isolates were selected from locations in Mexico to represent most years and regions with dengue activity and were collected between 1980 and 2002. The analyses of DEN-1 and DEN-4 isolates suggested there was a single introduction into the Americas in recent decades. A new DEN-3 Asian genotype was introduced into Central America in 1994 that was different from the strain associated with DHF outbreaks. Molecular epidemiological techniques are essential for understanding epidemic potential, trafficking, and evolutionary history of DEN virus.

**4. Molecular epidemiology of *Bunyaviruses***

**a. Crimean Congo hemorrhagic fever virus**

Phylogenetic analysis of 13 geographically and temporally diverse Crimean Congo hemorrhagic fever (CCHF) virus strains (*Nairovirus*) revealed highly significant

variability among the viruses (Deyde et al., 2006). Distinct geographic lineages exist, but with multiple exceptions indicative of long-distance virus movement. The analysis revealed that viruses segregated into distinct genotypes: group I, West Africa 1; II, Democratic of Congo; III, South Africa and West Africa 2; IV, Asia and the Middle East; V, Russia and Turkey; and VI, Greece. These groupings demonstrate that specific CCHF lineages move over large geographic distances.

**b. Rift Valley Fever virus**

A phylogenetic analysis of Rift Valley Fever virus (RVF, Genus: *Phlebovirus*) provided insight into the evolution and ecology of the virus (Bird et al., 2007). A complete genome sequence analysis of 33 ecologically and biologically diverse RVF strains demonstrated widespread virus movement, which suggests that the viral ecology of RVF is dynamic. Multiple introductions of virus genotypes may have occurred as these viruses have evolved from their most recent common ancestor.

**c. Akabane virus**

The S segment of the Akabane virus (*Orthobunyavirus*, Simbu serogroup) was characterized to define the genetic heterogeneity and molecular epidemiology of the virus (Yamakawa et al., 2006). The nucleotide and amino acid sequences were highly conserved among the isolates irrespective of the year of isolation and geographic origin.

**d. Inkoo and Tahyna viruses**

An analysis using the S segment of the Inkoo and Tahyna viruses (*Orthobunyavirus*, California serogroup) revealed that Inkoo virus is closely related to Jamestown Canyon virus, isolated in the United States, and Tahyna virus is genetically closest to Lumbo virus, isolated in Mozambique (Vapalahti et al., 1996). The data

suggested that genetic variation within the California serogroup is less related to geographic distance than to similarity in ecological niches.

**e. Jamestown Canyon virus**

The phylogenetic relationships of 56 Jamestown Canyon viruses (JCV, *Orthobunyavirus*, California serogroup) isolated from mosquitoes in Connecticut were determined for a 40-year period to evaluate the evolutionary pattern and genetic diversity of the viruses (Armstrong & Andreadis, 2007). Two major lineages were identified in Connecticut and viruses representing each lineage infected a diverse group of mosquito species over multiple years of sampling. The distribution of the lineages overlapped geographically in the collection area suggesting that the lineages are cocirculating in the same collection sites. There was a slight geographical structure along an east-west axis divided by the Connecticut River. The lineages identified in Connecticut were distinct from reference strains of JCV, Jerry Slough virus, South River virus, and Inkoo virus isolated from Colorado, California, New Jersey and Finland respectively, suggesting that there is genotypic variation observed within the collection region and amongst a broader geographic region (Armstrong & Andreadis, 2007).

**f. La Crosse virus**

LACV exhibits considerable evolutionary potential in nature. There are distinct geographic genotypes of the virus in different areas of the United States and there is some indication that disease severity may be conditioned by certain LACV genotypes (Armstrong & Andreadis, 2006, Huang et al., 1995, Huang et al., 1997, Klimas et al., 1981).

The genetic structure of LACV has been explored chronologically, ecologically, and geographically (El Said et al., 1979, Klimas et al., 1981). The genotypic variation of LACV isolated from different ecological niches and geographic regions of the United States was examined through the comparison of oligonucleotide fingerprints (Klimas et al., 1981). Viruses were isolated from the Midwest region of the United States (Wisconsin, Minnesota, Illinois, Indiana, and Ohio) and the eastern and southern portion of the United States (Texas, New York, Georgia, and North Carolina). Three haplotypes were observed: type A (Wisconsin, Minnesota, Indiana and Ohio) type B (Minnesota, Wisconsin and Illinois), and type C (eastern and southern region of the United States). An additional study compared LACV obtained from various ecological niches (i.e. a variety of mosquito species) in the northern United States and compared to the prototype LACV isolated from the first fatal case in 1960 (El Said et al., 1979, Thompson et al., 1965). Both studies found that no two viruses yielded identical fingerprints, suggesting there has been considerable evolution of the virus in nature. There is strong evidence that genetic drift occurring in nature is a factor in the genotypic variation of LACV isolates.

LACV isolates from mosquitoes and humans have also been analyzed using more sensitive techniques of sequencing and phylogenetic analysis to investigate genetic variation over time and distance (Huang et al., 1995, Huang et al., 1997). The M segment of LACV isolates from the brains of two children autopsied 18 years apart in Wisconsin were compared (Huang et al., 1995). There was an overall similarity of 99.6% in the nucleotide sequence. An additional phylogenetic study of LACV included an isolate recovered from a fatal case of encephalitis that occurred in Missouri in 1993, along with the two isolates from human fatalities in Wisconsin and multiple mosquito LACV

isolates (Huang et al., 1997). This allowed an opportunity to obtain sequence data from viruses that were separated spatially (upper Midwest and Missouri) and chronologically (33 years). The comparison of the M RNAs of these viruses showed human isolates had highly conserved nucleotide and deduced amino acid sequences. The high degree of conservation over time and space led to the hypothesis that human infections with this genotype of LACV is most likely to result in severe disease and a fatal outcome. Alternatively LACV passage in and isolation from the central nervous system of humans could select for specific virus genotypes, rendering false the hypothesis that fatal human infections are associated with a narrow range of genotypes. It is noteworthy in this regard that the M RNA segment sequence of virus isolated from mosquitoes was similar to that of the viruses recovered from humans suggesting that the genotype associated with human fatalities also circulates in mosquitoes (Huang et al., 1997).

An additional phylogenetic analysis of LACV M segment revealed three different lineages (Armstrong & Andreadis, 2006). One lineage consists of viruses isolated from the upper Midwest, one lineage is of viruses isolated from the eastern and southern region of the United States, and the third is a new lineage with a virus isolated from Connecticut. These studies support the previous observations found by Klimas *et al.* that there is considerable genetic diversity in LACV and that the viruses segregate into different lineages based upon a geographical distance (Klimas et al., 1981).

## H. RATIONALE AND OUTLINE OF DISSERTATION

In this dissertation the evolutionary, epidemic and maintenance potentials of LACV in nature were investigated. Four specific studies addressed these issues. The first study investigated the evolutionary potential of LACV through genetic shift. Molecular

epidemiological techniques were used to determine the frequency of segment reassortment of LACV in mosquitoes collected from the field (Chapter 2). The second study investigated the evolutionary potential of LACV by characterizing the gene flow of the viruses in the study area. Genetic variation of LACV isolates in the study area was investigated in terms of geographic origin, environmental terrain of collection site, and collection year (Chapter 3). The third study investigated the hypothesis that stabilized infection of a very small percentage of mosquitoes maintains LACV in nature. Field collected mosquitoes were assayed for LACV infection, super-infected mosquitoes were identified, and their viruses were characterized by sequences of genome segments (Chapter 4). The fourth study investigated a potential mechanism to enhance the maintenance and transmission of LACV in nature. The potential for LACV infection to increase insemination rates of TOT-LACV infected mosquitoes, thereby amplifying the number of infected progeny, was investigated (Chapter 5). The results of these studies provide insight to the evolutionary, epidemic, and maintenance potential of LACV.

# II. POTENTIAL FOR SEGMENT REASSORTMENT OF LA CROSSE VIRUS IN NATURE

## A. INTRODUCTION

In the 1970s, La Crosse virus (LACV family *Bunyaviridae*, genus *Orthobunyavirus*) emerged as a significant human pathogen in the upper Midwest, and is now the most common cause of pediatric arboviral encephalitis in the United States (Rust et al., 1999). LACV is primarily maintained in cycles between small mammals (usually chipmunks and tree squirrels) and *Aedes* (*Ochlerotatus*) *triseriatus* mosquitoes. *Aedes triseriatus* develop a life-long infection, and infected females can transovarially transmit (TOT) the virus to their progeny (Borucki et al., 2002, Watts et al., 1973). TOT is perhaps the most important mechanism for maintenance and amplification of LACV in nature (Beaty et al., 2000).

LACV has a tripartite, negative-sense RNA genome with the three segments designated large (L), medium (M), and small (S). The L segment encodes the RNA-dependent RNA polymerase (Endres et al., 1989), the M segment encodes a precursor polypeptide that is post-translationally cleaved to generate the G1 and G2 glycoproteins and a nonstructural protein, NSm (Elliott, 1985, Fuller & Bishop, 1982, Gentsch & Bishop, 1979), and the S segment encodes the nucleocapsid protein and a small nonstructural protein, NSs, in overlapping reading frames (Fuller & Bishop, 1982).

LACV exhibits considerable evolutionary potential in nature. There are distinct geographic genotypes of the virus in different areas of the United States (Armstrong & Andreadis, 2006, Huang et al., 1995, Huang et al., 1997, Klimas et al., 1981), and there is some indication that disease severity may be conditioned by certain LACV genotypes (Huang et al., 1997). The evolutionary success of the viruses in the family *Bunyaviridae*

is attributed in part to their ability to adapt to different conditions through genetic drift (intramolecular genetic changes) and genetic shift (segment reassortment).

Genetic drift occurs during genome replication and can result in viral diversity and altered fitness. RNA virus replication yields multiple genetic variants, or quasispecies, which occur due to the lack of proofreading enzymes and poor fidelity of the RNA polymerases. The error-prone polymerase can provide an array of mutations which allows constant adaptation to and selection by changes in the vector and vertebrate host, although a quasispecies with a greater fitness in one generation may not have greater fitness in future generations (Black & Salman, 2005).

Genetic shift (segment reassortment) can occur in mosquitoes that have become dually infected by ingesting viruses of two different LACV genotypes, either simultaneously or within two days of each other (Beaty et al., 1985). LACV reassortant viruses can be isolated from up to 25% of dually infected *Ae. triseriatus* mosquitoes, and the newly generated viruses can be transmitted (Beaty et al., 1985). The potential for segment reassortment increases when a transovarially-infected mosquito takes a blood meal from a viremic host (Borucki et al., 1999). These mosquitoes can be orally super-infected, and these dually infected mosquitoes can transmit new reassortant viruses (Borucki et al., 1999). The new reassortant viruses might exhibit new characteristics such as altered host and vector ranges, new tropisms or virulence, and thus may be epidemiologically significant (Beaty & Calisher, 1991). Segment reassortment is apparently restricted to closely related bunyaviruses, typically in the same serogroup (Chandler et al., 1990, Chandler et al., 1991, Gentsch et al., 1977, Urquidi & Bishop, 1992)

There is evidence for reassortment between LACV genotypes in nature. For example, the genomes of 23 isolates of LACV were analyzed by oligonucleotide fingerprinting and categorized in terms of the degree of their RNA sequence relatedness. One genotype (denoted type A) was isolated from mosquitoes from Wisconsin, Minnesota, Indiana, and Ohio and a second genotype (denoted type B) was isolated from mosquitoes from Minnesota, Wisconsin, and Illinois. A reassortant LACV isolated in Rochester, MN contained the S segment of the B genotype, and the M and L segments of the A genotype (Klimas et al., 1981).

There is also evidence for genome reassortment among other *Orthobunyaviruses*. Ngari virus is a newly emerged reassortant virus associated with severe disease epidemics in Africa (Gerrard et al., 2004). Sequence analysis of the three genomic RNA segments revealed that the S and L segments were derived from Bunyamwera virus, but the M segment was derived from the Batai virus, an *Orthobunyavirus* that was first detected in Malaysia (Briese et al., 2006). Group C *Orthobunyaviruses* also reassort (Nunes et al., 2005). Phylogenetic analysis revealed that Caraparu virus contains an S segment sequence that is nearly identical to that of the Oriboca virus and therefore is a natural reassortant virus.

Reassortment of viruses in other *Bunyaviridae* genera has also been documented. Reassortant Sin Nombre viruses (Hantavirus) have been detected in rodents in nature (Li et al., 1995a) and reassortant Crimean Congo hemorrhagic fever viruses (Nairovirus) (Hewson et al., 2004) have also been detected. Clearly genome reassortment of viruses in the family *Bunyaviridae* is widespread in nature. The epidemiologic consequences of these evolutionary events are poorly understood.

Segment reassortment has been thought to occur frequently in the *Bunyaviridae* family. In this study through the use of sensitive molecular epidemiological techniques, the evolutionary and reassortment potential of LACV in field-infected mosquitoes from the upper Midwest of the United States was investigated.

## B. MATERIALS AND METHODS

### 1. Egg collection

*Aedes triseriatus* eggs were collected from five oviposition traps in each of 151 sites in Minnesota (n = 37), Wisconsin (n = 108) and Iowa (n = 6) in areas where LACV encephalitis cases occurred or areas that contained clusters of people judged by the La Crosse County Public Health Department to be at risk for infection (e.g. wooded areas adjacent to houses with children, schools, or playgrounds). Mosquito eggs that had entered diapause in fall 2000 were collected in the spring of 2001. Mosquito eggs were also collected between mid-June and August of 2004. Eggs were collected in Crawford, La Crosse, Monroe, Vernon and Iowa counties in Wisconsin; Winona, Houston, and Grant counties in Minnesota; and Clayton and Allamakee counties in Iowa (Figure 2.1). The eggs were transported to the insectaries at the Arthropod-borne and Infectious Diseases Laboratory (AIDL) at Colorado State University (CSU), Fort Collins, CO. The eggs collected in 2001 and 2004 were immediately hatched, and progeny were reared.

### 2. LACV analytes

Previously isolated LACV strains were also used in the analysis (Table 2.1). A LACV isolate recovered in 1970 is genotype A and it was passed five times in suckling mouse brains (SMB), once in *Ae. triseriatus* mosquitoes and twice in baby hamster kidney (BHK-21) cell culture. A 1978 LACV (78V-8853) was isolated from

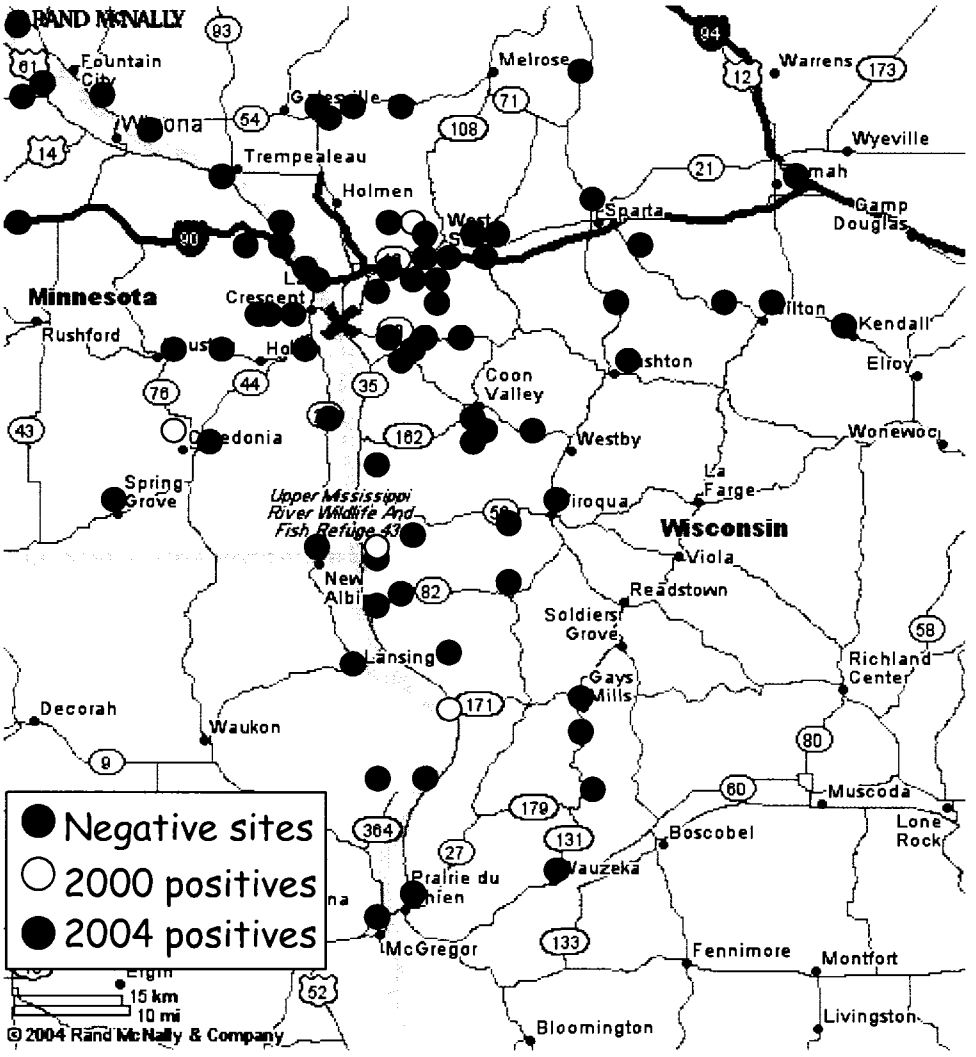



**Figure 2.1: Mosquito collection sites in Minnesota, Wisconsin, and Iowa**
Circles represent all collections sites. Yellow circles are the sites where LACV positive mosquitoes were collected in 2000, red circles are the sites where LACV positive mosquitoes were collected in 2004, and black circles were the sites without positive mosquitoes.

an *Ae. triseriatus* mosquito from Rochester, MN and passed once in Vero cells and twice in SMB. A LACV isolate recovered in 1979 had been passed four times in BHK-21 cells. The passage history for the 1981 LACV isolate is unknown.

**Table 2.1: LACV isolates analyzed in the segment reassortant analysis**

| | | | |
|---|---|---|---|
| 1970 | | | Suckling mouse 5, *Ae. triseratus* 1, BHK-21 2 |
| LAC 78V-8853 (1978) | *Ae. triseriatus* | Rochester, MN | Vero 1, Suckling mouse 2 |
| 1979 | | | BHK-21 4 |
| 1981 | | | unknown |

### 3. Immunofluorescence assay (IFA)

To determine if mosquitoes were infected, mosquito heads were severed, squashed onto acid-washed microscope slides, and fixed in acetone. They were assayed for LACV antigen by direct immunofluorescence assay (IFA) using a LACV-specific polyclonal antibody (Beaty & Thompson, 1975).

### 4. RNA purification amplification by reverse transcription-PCR

The posterior half of each mosquito abdomen was individually homogenized in 500µl of Trizol (Invitrogen), using a pellet pestle (Fisher Scientific), then total RNA was extracted according to manufacturer's instructions. The viruses from 1970, 1978, 1979, and 1981 were prepared as follows. The 1979 isolate was intrathoracically inoculated (~0.4µl each) into 10 mosquitoes. Fourteen days post-injection, the mosquitoes were tested for LACV antigen by IFA. Each positive mosquito was triturated in 500µl of Trizol and RNA was purified according to manufacturer's instructions. The RNA from the 1978, 1980 and 1981 isolates was extracted from C6 36 cell cultures infected at a multiplicity of infection (MOI) of 0.01. Three days post-infection, cells were scraped into the medium, centrifuged and the cell pellet was resuspended in 500µl of Trizol for RNA recovery according to manufacturer's instructions.

Portions of the LACV S, M, and L RNA segments were transcribed to cDNA using Superscript II reverse transcriptase (Invitrogen) and amplified by PCR using Ex Taq DNA polymerase (Takara) according to manufacturer's instructions. The primers specific for the S segment (forward: 5'-GCAAATGGATTTGATCCTGATGCAG-3', reverse: 5'-CTTAAGGCCTTCTTCAGG TATTGAG-3') amplified a 461 nucleotide region (nucleotides 144 to 605) of the nucleocapsid gene and the nonstructural protein that was selected because it was the most variable region of the published S sequences. The S segment is 984 nucleotides in length, so the amplified region encompasses almost half the entire segment. The primers specific for the M segment (forward: 5'-CCAAAAGCAACA AAAGAAAGA-3', reverse: 5'- CTGAAGGCATGA TGCAAAG-3') amplified a highly variable 410 nucleotide region in the 5' half of the G1 gene, (nucleotides 1585 to 1995) (Borucki et al., 2001). The primers specific for the L segment (forward: 5'-GCATGTGTAGCCA AGGATATCGATG-3', reverse: 5'-CAGTCTTGCACCAGGGTG CTGTAAG-3') amplified a 486 nucleotide region (nucleotides 140 to 626). These primers also were selected to amplify the most variable region of the L segment. Primers specific for the *Ae. triseriatus* ribosomal protein Rpl34 mRNA were used as a positive control. PCR was performed as follows: 94°C for 5 minutes, 35 cycles of [94°C for 1 minute, 55°C for 1 minute and 72°C for 1 minute] followed by a final extension at 72°C for 8 minutes.

**5. Amplicon cloning and sequencing**

PCR products were separated by electrophoresis in 1% agarose gels with TAE buffer, stained with ethidium bromide, excised and extracted using a kit from Marligen Biosciences according to manufacturer's instructions. PCR products were inserted into

the pCR4-TOPO cloning vector (Invitrogen) and resulting plasmids were used to transform competent TOP10 *E. coli* cells (Invitrogen). Cells were grown on LB agar containing ampicillin (50μg/ml) and kanamycin (50μg/ml). Colonies were screened for inserts by PCR amplification, using the original primers and positive products were purified using a QIAquick spin column (Qiagen). Three to five cDNA clones per segment from each mosquito were sequenced using the ABI PRISM dye terminator cycle sequencing kit (Applied Biosystems, CA) and the ABI 310 DNA automated sequencer at Macromolecular Resources, CSU. A 415 nucleotide region of the S segment (nucleotides 190-604), a 358 nucleotide region of the M segment (nucleotides 1637-1994), and a 447 nucleotide region of the L segment (nucleotides 179-625) were sequenced.

**6. Haplotype determination**

Genetic groups were established for each of the three segments through maximum parsimony analysis, sequence identity matrix, neighbor joining distance matrix, and ratio of synonymous to nonsynonymous substitutions. Two different grouping systems were utilized: (1) a system in which the haplotypes were determined by the original phylogenetic analysis and (2) a conservative grouping system.

**7. Statistical analyses**

**a. Linkage Disequilibrium Analysis**

A linkage disequilibrium analysis was performed to determine if the segments assort randomly, thereby suggesting segment reassortment. Linkage disequilibrium describes a situation in which some combinations of segments occur together more or less frequently in a mosquito than would be predicted by their independent frequencies.

The first step is to determine the number of times segment, $S_i$,$M_j$, and $L_k$ appear in the same mosquito, where for example, $S_i$ is a unique sequence of the segment.

$T_{ijk}$ = the number of times haplotypes *i, j,* and *k* occur in a mosquito. (1)

$E_{ijk}$ is the number of times haplotypes *i, j,* and *k* are expected to occur in a mosquito.

$$E_{ijk}= (p_i \times p_j \times p_k) \times N, \quad (2)$$

where $p_i$ is the frequency of $S_i$ in the mosquito population and N is the number of mosquitoes.

Linkage disequilibrium ($D_{ijk}$) was then estimated.

$$D_{ij} = (N/N-1)*( T_{ijk} - E_{ijk})/N \quad (3)$$

A correlation coefficient $R_{ijk}$ was determined (Hill & Robertson, 1966).

$$R_{ijk} = D_{ij}/((p_i(1-p_i))(p_j(1-p_j))(p_k(1-p_k))^{1/2} \quad (4)$$

A chi-square statistic ($\chi^2_{Link}$) and the corresponding level of significance were calculated for each group of haplotypes to test the hypothesis that the individual haplotype combinations are in linkage equilibrium.

$$\chi^2_{Link\,(1d.f)}= (N R_{ijk}^2) \quad (5)$$

An additional chi-square statistic for independence ($\chi^2_{Ind}$) was calculated to determine if there was significant dependence for the interaction between the three segments overall.

$$\chi^2_{Ind} = (T_{ijk} - E_{ijk})^2/E_{ijk} \quad (6)$$

**b. Maximum Parsimony analysis**

Maximum parsimony phylogenetic analysis was performed using the Phylogenetic Analysis Using Parsimony (PAUP) 4.0b10 package (Swofford, 2003). The phylogenetic trees indicate the branches that appeared in the majority of the 1000 bootstrap pseudo replications and the frequency with which these appear among

replications. A maximum parsimony phylogenetic tree was created for each of the three genome segments.

## C. RESULTS

### 1. LACV infected mosquitoes analyzed

We investigated the genetic diversity and segment reassortment potential of LACV in the upper Midwestern U.S. A total of 6,791 mosquitoes collected as eggs at study sites in Wisconsin, Minnesota, and Iowa (Figure 2.1) were reared and tested for LACV antigen by IFA. Of these, 309 (4.5%) were positive. Viral RNA was amplified from one to three mosquitoes from the selected sites listed in Table 2.2. Viral RNA from 40 mosquitoes was analyzed, including 34 field collected mosquitoes from 2004 and six field collected mosquitoes from 2000. In addition, RNA from four virus isolates from 1970, 1978, 1979, and 1981, was analyzed (Table 2.1).

### 2. Sequence polymorphism and nucleotide substitution

Genome sequence polymorphism (theta) in a population is the proportion of nucleotide sites that are expected to be polymorphic in any sample in a particular region of the genome. The value takes into account the number of nucleotide differences between each pair of sequences and the number of possible pairs (Kimura, 1969). Genome sequence polymorphism for the L, M, and S segment sequences were 0.00827, 0.00985, and 0.00661, respectively. This can be seen graphically in Figure 2.2. The polymorphic sites of the L segment sequences can be observed between nucleotides 480 and 600. The polymorphic sites of the M segment sequences are found between nucleotides 1600 and 1800, and nucleotides 1900 and 1985. There are not as many

**Table 2.2: Field collected *Ae. triseriatus* mosquitoes from Minnesota and Wisconsin in 2004 and 2000 that were tested for LACV infection*.**

| 2004 samples | | | | | |
|---|---|---|---|---|---|
| BCA/Winona/2004 | Winona, MN | 6/18/2004 | 7 | 3 | 42.9% |
| BCB/Winona/2004 | Winona, MN | 6/18/2004 | 7 | 3 | 42.9% |
| BRS/Houston/2004 | Houston, MN | 6/29/2004 | 50 | 1 | 2.0% |
| BWL/Houston/2004 | Houston, MN | 6/28/2004 | 38 | 2 | 5.3% |
| CAL-B/Houston/2004 | Houston, MN | 7/20/2004 | 50 | 2 | 4.0% |
| CAL-D/Houston/2004 | Houston, MN | 7/20/2004 | 50 | 5 | 10.0% |
| CAL-GA/Houston/2004 | Houston, MN | 7/20/2004 | 50 | 5 | 10.0% |
| CAT/Monroe/2004 | Monroe, WI | 7/19/2004 | 50 | 1 | 2.0% |
| DAK-90/Winona/2004 | Winona, MN | 6/18/2004 | 38 | 3 | 7.9% |
| ESO/Vernon/2004 | Vernon, WI | 7/22/2004 | 50 | 4 | 8.0% |
| GAY120/Crawford/2004 | Crawford, WI | 7/22/2004 | 50 | 12 | 24.0% |
| GRL/La Crosse/2004 | La Crosse, WI | 7/19/2004 | 50 | 1 | 2.0% |
| HCS/Houston/2004 | Houston, MN | 8/2/2004 | 42 | 1 | 2.4% |
| HHS/Houston/2004 | Houston, MN | 7/2/2004 | 50 | 1 | 2.0% |
| H0/Vernon/2004 | Vernon, WI | 6/21/2004 | 24 | 1 | 4.2% |
| INNSL/La Crosse/2004 | La Crosse, WI | 6/28/2004 | 20 | 3 | 15.0% |
| LAXCC/La Crosse/2004 | La Crosse, WI | 6/28/2004 | 30 | 2 | 6.7% |
| MCBB/La Crosse/2004 | La Crosse, WI | 7/8/2004 | 50 | 2 | 4.0% |
| MCPA/La Crosse/2004 | La Crosse, WI | 6/17/2004 | 35 | 2 | 5.7% |
| MCPB/La Crosse/2004 | La Crosse, WI | 6/17/2004 | 35 | 2 | 5.7% |
| NAT/Crawford/2004 | Crawford, WI | 7/12/2004 | 50 | 3 | 6.0% |
| NFCS/Winona/2004 | Winona, MN | 7/19/2004 | 50 | 1 | 2.0% |
| OTS/La Crosse/2004 | La Crosse, WI | 7/19/2004 | 50 | 3 | 6.0% |
| RRA/Houston/2004 | Houston, MN | 7/12/2004 | 50 | 3 | 6.0% |
| RCS/Crawford/2004 | Crawford, WI | 6/21/2004 | 50 | 6 | 12.0% |
| SHR/Vernon/2004 | Vernon, WI | 6/21/2004 | 50 | 1 | 2.0% |
| SRS/La Crosse/2004 | La Crosse, WI | 7/19/2004 | 50 | 7 | 14.0% |
| SST/La Crosse/2004 | La Crosse, WI | 7/19/04 | 50 | 2 | 4.0% |
| SVP/Vernon/2004 | Vernon, WI | 7/26/04 | 50 | 4 | 8.0% |
| TFP/Monroe/2004 | Monroe, WI | 7/19/04 | 41 | 17 | 41.0% |
| VSB/Vernon/2004 | Vernon, WI | 6/21/04 | 50 | 5 | 10.0% |
| WKCS/Crawford/2004 | Crawford, WI | 6/21/04 | 27 | 3 | 11.1% |
| WSB/La Crosse/2004 | La Crosse,WI | 7/20/04 | 47 | 2 | 4.3% |
| WBF/Monroe/2004 | Monroe, WI | 7/19/04 | 50 | 8 | 16.0% |
| **2000 Samples** | | | | | |
| CAL-BA/Houston/2000 | Houston, MN | 5/1/2001 | 50 | 5 | 10.0% |
| CAL-BB/Houston/2000 | Houston, MN | 5/1/2001 | 50 | 5 | 10.0% |
| INNB/La Crosse/2000 | La Crosse, WI | 5/1/2001 | 50 | 2 | 4.0% |
| LRHE-A/La Crosse/2000 | La Crosse, WI | 5/1/2001 | 50 | 2 | 4.0% |
| LRHE-B/La Crosse/2000 | La Crosse, WI | 5/1/2001 | 50 | 2 | 4.0% |
| VSA/Vernon/2000 | Vernon, WI | 5/1/2001 | 50 | 1 | 2.0% |

*Fifty mosquitoes were tested for LACV antigen from most sites. There were 13 sites with less than 50 adult mosquitoes.

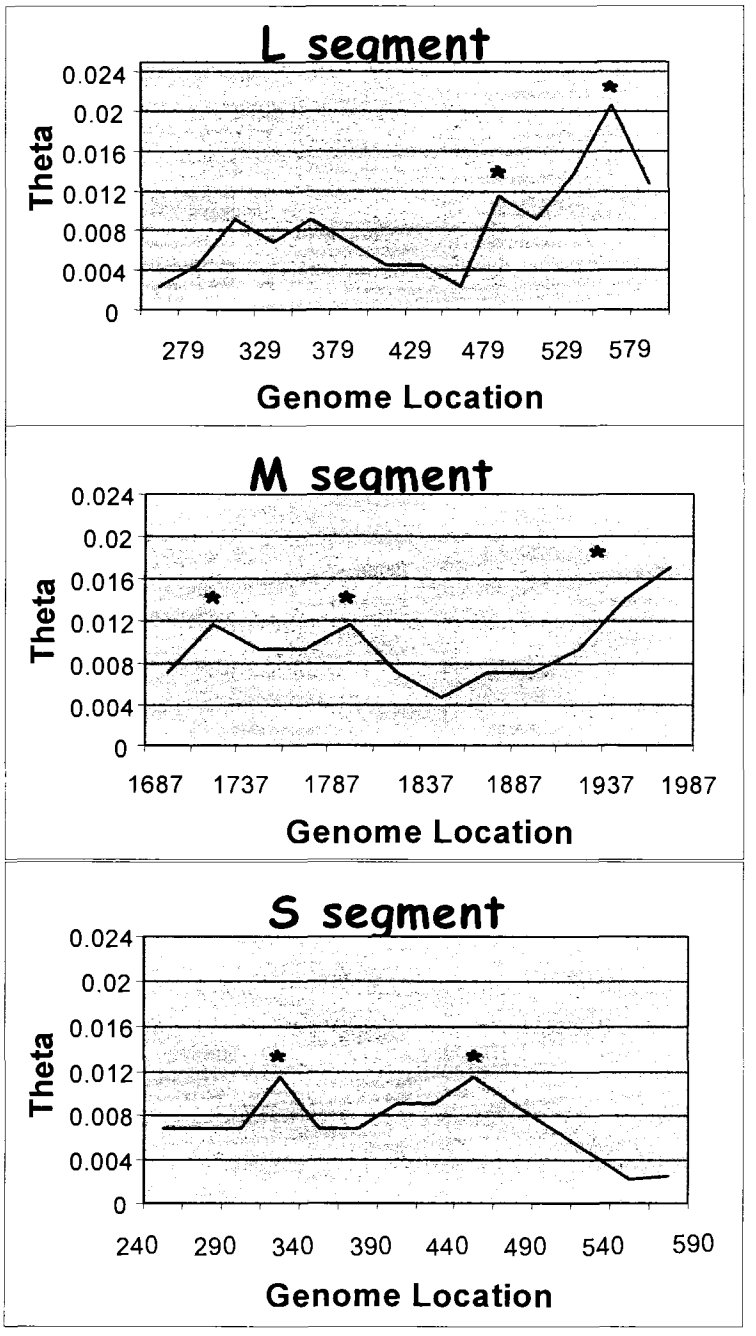



**Figure 2.2: Nucleotide polymorphism (theta) of the LACV L, M and S segment sequences amplified from field-infected mosquitoes**

Polymorphic regions are regions of 50 nucleotides with 5 or more polymorphisms and are represented by the asterisks. The polymorphic regions of the L segment sequences can be observed between nucleotides 480 and 600. The polymorphic regions of the M segment sequences are found between 1600 and 1800, and 1900 and 1985. The polymorphic regions observed in the S segment sequences are found between 300 and 350, and between 450 and 500.

polymorphic sites observed in the S segment sequences; they are found between nucleotides 300 and 350 and between nucleotides 450 and 500.

All three segments had more nonsynonymous substitutions than synonymous, suggesting that the nonsynonymous mutations are maintained by positive selection. The M segment had four synonymous substitutions and ten nonsynonymous substitutions. The L segment had three synonymous substitutions and thirteen nonsynonymous substitutions. The S segment had three synonymous substitutions and nine nonsynonymous substitutions. The nonsynonymous substitutions that were found are shown in Table 2.3.

**3. Haplotype determination**

Two different haplotype grouping systems were developed: (1) a system in which the haplotypes were determined by the original phylogenetic analysis and (2) a conservative grouping system. The original system identified four S clades based on seven polymorphic sites, five of which were nonsynonymous mutations. The three clades identified in the M segment were based on ten polymorphic sites, five of which were nonsynonymous. For the L segment, four clades were identified based on six polymorphic sites, all of which were nonsynonymous substitutions (Figure 2.3). Because some of these clades differed by only a few nucleotides, the conservative grouping system was established, which reduced the number of clades to three S haplotypes, three M haplotypes and two L haplotypes (Figure 2.4).

**Table 2.3: Nonsynonymous mutations found in sequences of LACV RNA amplified from field collected mosquitoes***

| | | | |
|---|---|---|---|
| L | 252 | C → A | Pro → His |
| L | 282 | C → A | Pro → Glu |
| L | 313 | G → A | Met → Ile |
| L | 321 | T → C | Tyr → Ala |
| L | 374 | A → G | Thr → Ala |
| L | 489 | A → G | Asp → Gly |
| L | 490 | T → C | Arg → Gly |
| L | 536 | A → G | Asn → Asp |
| L | 547 | T → G | Phe → Cys |
| L | 555 | A → G | Lys → Arg |
| L | 561 | T → C | Ser → Leu |
| L | 576 | T → G | Phe → Cys |
| L | 608 | G → A | Ala → Thr |
| M | 1663 | A → G | Ile → Met |
| M | 1749 | G → A | Asn → Ser |
| M | 1754 | T → C | Tyr → His |
| M | 1782 | A → G | Asp → Gly |
| M | 1815 | T → C | Val → Ala |
| M | 1826 | T → C | Ser → Pro |
| M | 1898 | T → C | Cys → Arg |
| M | 1913 | T → C | Trp → Arg |
| M | 1961 | A → G | Lys → Glu |
| M | 1964 | T → C | Phe → Leu |
| S | 209 | T → C | Phe → Ser |
| S | 273 | A → C | Lys → Asn |
| S | 298 | A → G | Ile → Val |
| S | 340 | A → G | Asn → Asp |
| S | 347 | A → G | Asp → Gly |
| S | 400 | T → C | Tyr → His |
| S | 419 | A → T | Glu → Val |
| S | 445 | A → G | Thr → Ala |
| S | 463 | G → A | Ala → Thr |

*Sequence analysis revealed numerous nonsynonymous mutations for the three segments, L, M and S. There were thirteen nonsynonymous mutations observed in a 447 nucleotide region of the L segment, ten nonsynonymous mutations in a 358 nucleotide region of the M segment, and nine nonsynonymous mutations in a 415 nucleotide region of the S segment.

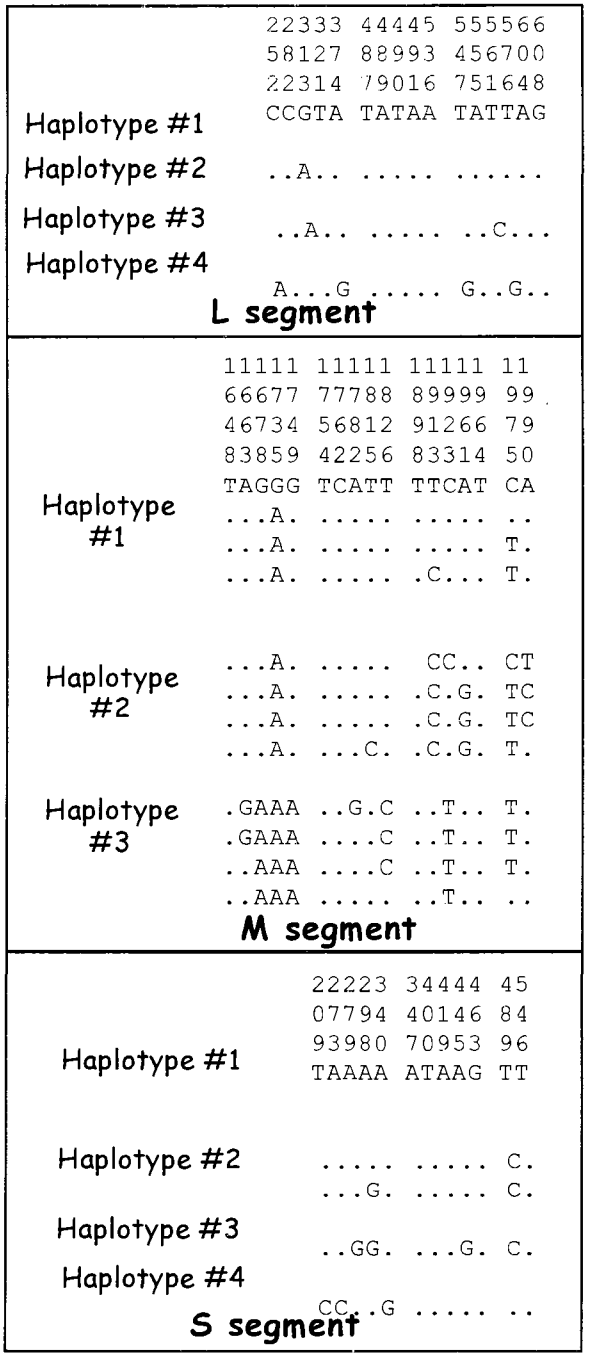



**Figure 2.3: LACV L, M, and S segment haplotype determination based on the original approach**

Phylogenetic analyses yielded four haplotypes for the L segment (A), three haplotypes for the M segment (B), and four haplotypes for the S segment (C). The genome position is provided above the genetic sequence.

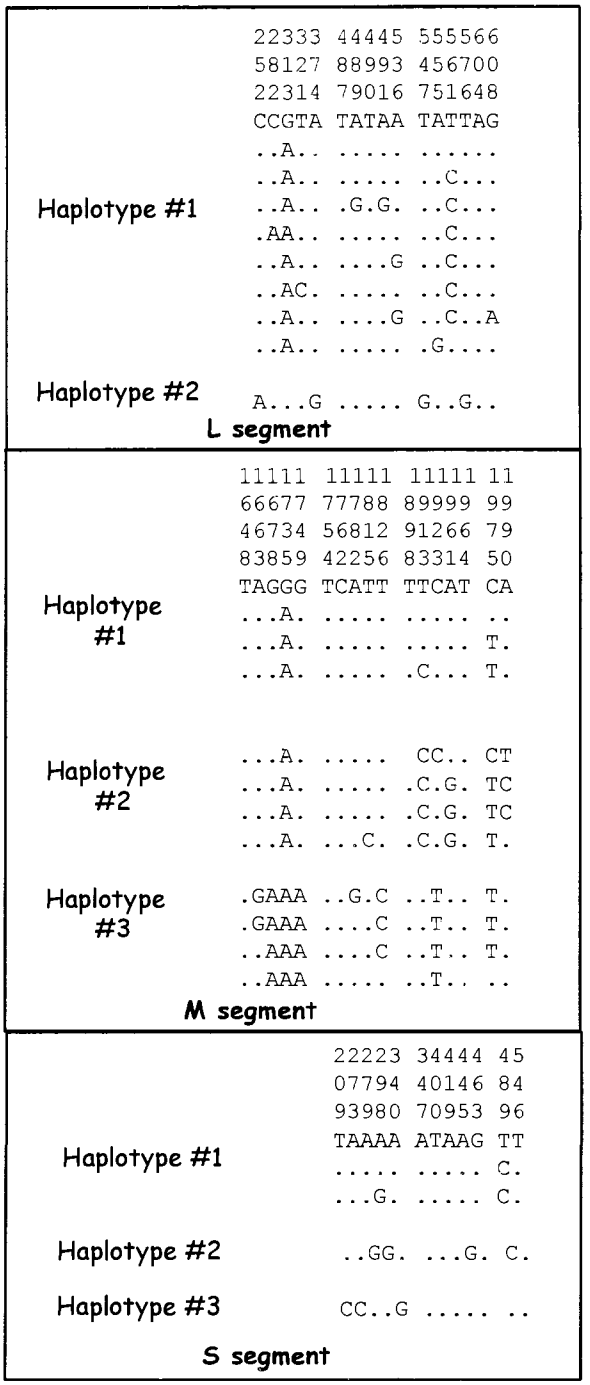



**Figure 2.4: LACV L, M, and S segment haplotype determination based on the conservative approach**

Conservative phylogenetic analyses yielded two haplotypes for the L segment (A), three haplotypes for the M segment (B), and three haplotypes for the S segment (C). The genome position is provided above the genetic sequence.

**4. Phylogenetic analysis**

Comparison of the clades on the three maximum parsimony trees provides evidence for the occurrence of segment reassortment (Figures 2.5-2.7). If reassortment had not occurred, the clades in each of the three trees would be identical. A number of mosquitoes contain segments that clustered into different clades in each of the trees. For example, the sample MCBB/La Crosse/2004 is in the group #2 for the L segment and in the group #3 for the S segment. Another example is the mosquito collected from NFCS/ Winona/2004. It is in group #3 in the L segment, group #2 in the M segment and group #4 in the S segment. Patterns such as this strongly suggest that segment reassortment is occurring.

**5. Linkage disequilibrium**

Linkage disequilibrium is a gauge of whether reassortment has occurred. Segments that are randomly assorted are considered to be in a state of linkage equilibrium and segments not randomly assorted are considered to be in linkage disequilibrium.

Linkage disequilibrium indicates that segments assort more or less often than expected by their independent frequencies and therefore segment reassortment has not occurred. This analysis was performed to determine if and how much segment reassortment was occurring in the field. Through the clades identified by phylogenetic analysis, it was determined that 17 of the 44 samples (38.6%) are in linkage equilibrium, suggesting that genetic variation occurring through segment reassortment is quite frequent in the field (Table 2.4). In the conservative analysis, 11 of the 44 samples (25.0%) were in linkage equilibrium, supporting the hypothesis that segment

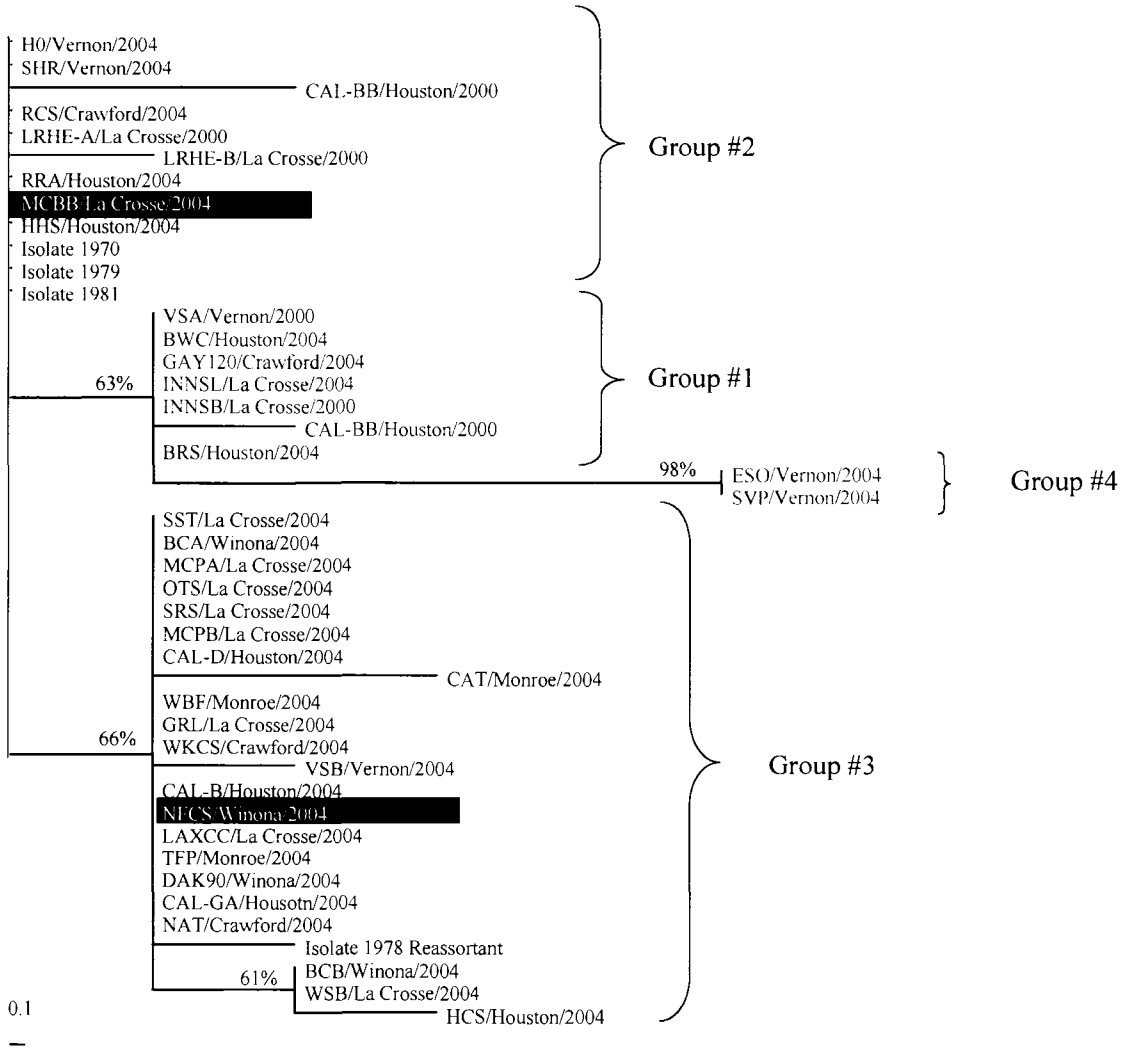
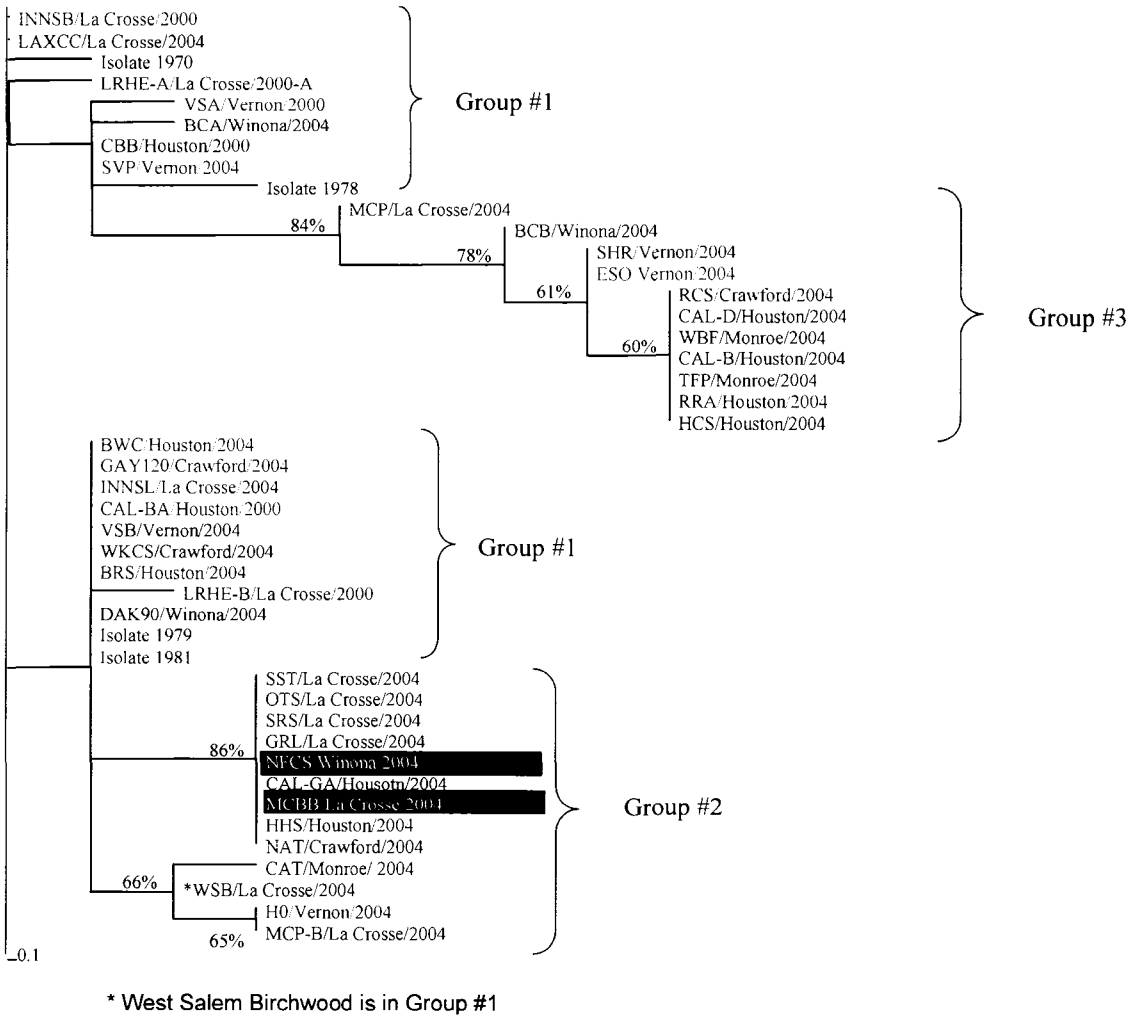



**Figure 2.5: Maximum parsimony phylogenetic analysis of the L segment of LACV RNA amplified from field collected mosquitoes**

Phylogenetic analysis of LACV sequence amplified from field-infected mosquitoes collected in 2000 and 2004 and from LACV virus isolates from 1970, 1978, 1979 and 1981 for the L segment from nucleotides 140-626. Bootstrap values were assigned for 100 replicates represented by the numbers on the branches. Colors represent genotypes determined for the L segment and are continued for the S and M segments. The two highlighted samples are examples of segment reassortment.




* West Salem Birchwood is in Group #1

**Figure 2.6: Maximum parsimony phylogenetic analysis of the M segment of LACV RNA amplified from field collected mosquitoes**

Phylogenetic analysis of LACV sequence amplified from field-infected mosquitoes collected in 2000 and 2004 from LACV virus isolates from 1970, 1978, 1979 and 1981 for the M segment from nucleotides 1585-1995. Bootstrap values were assigned for 100 replicates represented by the numbers on the branches. Colors represent genotypes determined for the L segment and are continued for the S and M segments. The two highlighted samples are examples of segment reassortment.

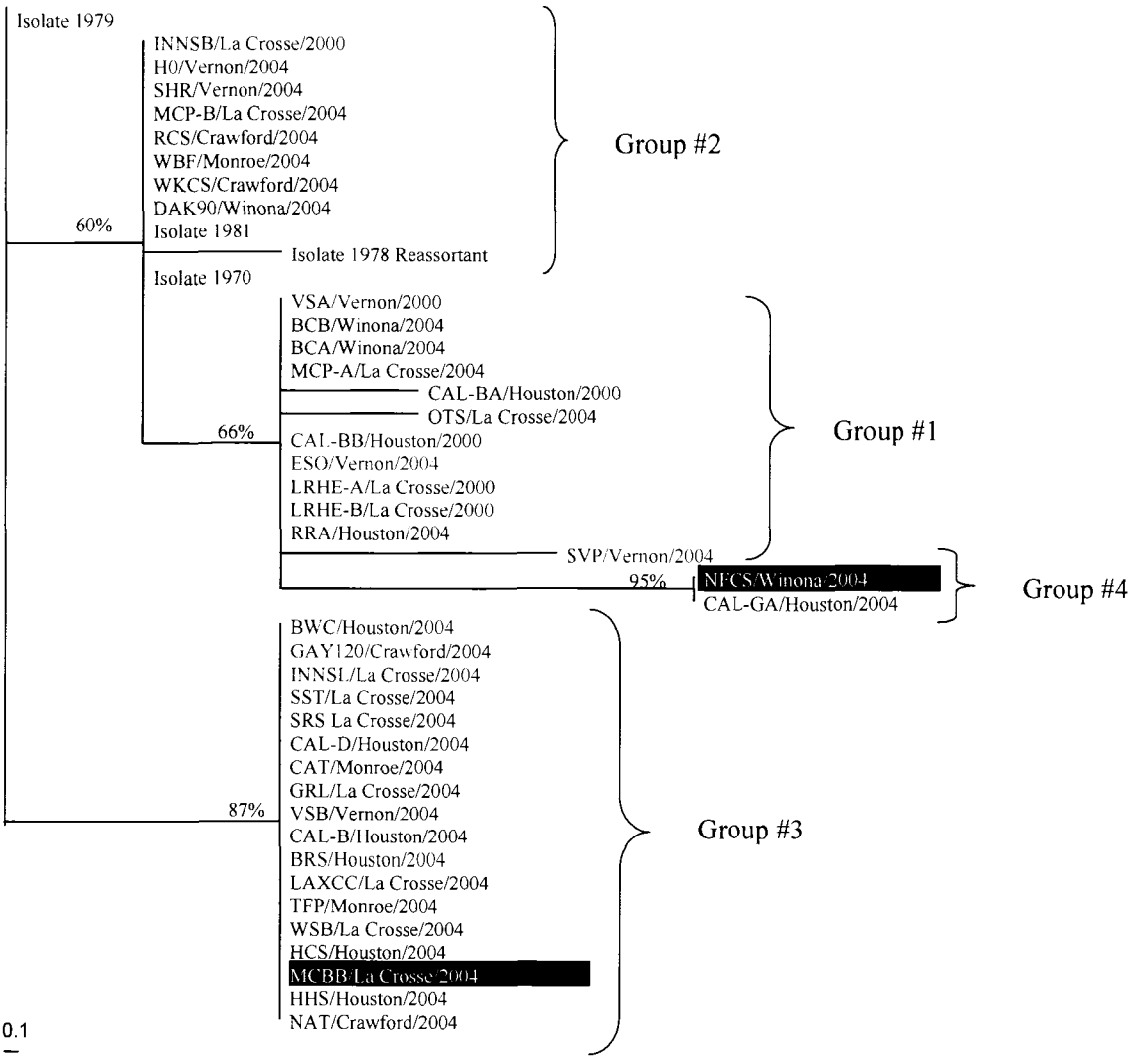



**Figure 2.7: Maximum parsimony phylogenetic analysis of the S segment of LACV RNA amplified from field collected mosquitoes**

Phylogenetic analysis of LACV sequence amplified from field-infected mosquitoes collected in 2000 and 2004 from LACV virus isolates from 1970, 1978, 1979 and 1981 for the S segment from nucleotides 144-605. Bootstrap values were assigned for 100 replicates represented by the numbers on the branches. Colors represent genotypes determined for the L segment and are continued for the S and M segments. The two highlighted samples are examples of segment reassortment.

reassortment is a significant mechanism of evolution for LACV (Table 2.5). A Chi-square test ($\chi^2_{Link}$) determines if the haplotypes combinations are in linkage disequilibrium. The null hypothesis is that the segments from each mosquito clustered independently. The haplotype combinations deemed significant through the $\chi^2_{Link}$ test are considered to be in linkage disequilibrium and therefore not segregating independently. In the analyses 25-38.6% of the samples yielded an insignificant p-value from the $\chi^2_{Link}$ test. These are therefore in linkage equilibrium and are possible reassortants (Tables 2.4 and 2.5).

An additional analysis was performed to determine which of the three segments were most likely to be randomly segregating. A two by two analysis was done comparing the possible haplotype combinations between S and M segments, the M and L segments and the S and L segments (Table 2.6). The S and M analysis revealed that 42 samples of the 44 (95.5%) are in linkage equilibrium using the $\chi^2_{Link}$ test. A similar analysis of the conservative clades revealed that 39 of the 44 samples (88.6%) are in equilibrium. In the M and L analysis, 29 samples of the 44 (65.9%) were revealed to be in a state of linkage equilibrium, and all the combinations were in linkage equilibrium in the conservative analysis. The S and L analysis yielded similar results where 36 of the 44 mosquito samples (81.8%) are in linkage equilibrium compared to all the combinations in the conservative approach (Table 2.7).

The $\chi^2_{Ind}$ test revealed significant relationships between the three segments, confirming that majority of the haplotype combinations are in linkage disequilibrium (Tables 2.4 and 2.5). The $\chi^2_{Ind}$ test was also performed to compare all possible haplotype combinations between two segments (S and M, M and L, and S and L), and the results

## Table 2.4: LACV segment reassortment occurs in field collected mosquitoes as revealed by a linkage disequilibrium analysis*

| | | | | | |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 2 | -0.31122 | * |
| **1** | **1** | **2** | **2** | **-0.13718** | |
| **1** | **1** | **3** | **2** | **0.23598** | |
| 1 | 1 | 4 | 1 | -0.37123 | * |
| 1 | 2 | 1 | 0 | 0.16690 | |
| 1 | 2 | 2 | 0 | 0.22154 | |
| **1** | **2** | **3** | **1** | **0.18487** | |
| 1 | 2 | 4 | 0 | 0.08373 | |
| 1 | 3 | 1 | 0 | 0.15736 | |
| **1** | **3** | **2** | **1** | **-0.06962** | |
| **1** | **3** | **3** | **2** | **-0.08807** | |
| 1 | 3 | 4 | 1 | -0.49999 | *** |
| **2** | **1** | **1** | **1** | **-0.02539** | |
| 2 | 1 | 2 | 3 | -0.37861 | * |
| **2** | **1** | **3** | **3** | **0.02603** | |
| 2 | 1 | 4 | 0 | 0.13066 | |
| 2 | 2 | 1 | 0 | 0.16690 | |
| **2** | **2** | **2** | **1** | **-0.04923** | |
| **2** | **2** | **3** | **1** | **0.18487** | |
| 2 | 2 | 4 | 0 | 0.08373 | |
| 2 | 3 | 1 | 0 | 0.15736 | |
| 2 | 3 | 2 | 2 | -0.34812 | * |
| **2** | **3** | **3** | **1** | **0.15412** | |
| 2 | 3 | 4 | 0 | 0.07895 | |
| 3 | 1 | 1 | 4 | -0.68177 | *** |
| 3 | 1 | 2 | 0 | 0.46970 | ** |
| 3 | 1 | 3 | 3 | 0.32064 | * |
| 3 | 1 | 4 | 0 | 0.17753 | |
| 3 | 2 | 1 | 0 | 0.39821 | ** |
| **3** | **2** | **2** | **2** | **-0.18953** | |
| 3 | 2 | 3 | 5 | -0.49534 | ** |
| 3 | 2 | 4 | 0 | 0.11377 | |
| 3 | 3 | 1 | 0 | 0.21381 | |
| 3 | 3 | 2 | 0 | 0.28380 | |
| 3 | 3 | 3 | 4 | -0.33904 | * |
| 3 | 3 | 4 | 0 | 0.10727 | |
| 4 | 1 | 1 | 0 | 0.09280 | |
| 4 | 1 | 2 | 0 | 0.12319 | |
| 4 | 1 | 3 | 0 | 0.23373 | |
| 4 | 1 | 4 | 0 | 0.04656 | |
| 4 | 2 | 1 | 0 | 0.05948 | |
| 4 | 2 | 2 | 0 | 0.07895 | |
| 4 | 2 | 3 | 2 | -0.85713 | *** |
| 4 | 2 | 4 | 0 | 0.02984 | |
| 4 | 3 | 1 | 0 | 0.05607 | |
| 4 | 3 | 2 | 0 | 0.07443 | |
| 4 | 3 | 3 | 0 | 0.14122 | |
| 4 | 3 | 4 | 0 | 0.02813 | |
| | | | 44 | | |
| **$X^2_{Ind.}$ Test** | 44.15369 | | | | |
| **p-value** | 0.00055 | 18 d.f. | | | |

*The 48 possible haplotype combinations were analyzed for linkage disequilibrium with four S segments, three M segments and four L segments. The observed column is the number of that specific haplotype combination found in the population. The $R_{ijk}$ is the correlation coefficient for linkage disequilibrium. The highlighted samples are those haplotype combinations that are in linkage equilibrium that were observed in the population.
* p-value $\leq 0.05$, ** p-value $\leq 0.01$, *** p-value $\leq 0.001$

**Table 2.5: LACV segment reassortment occurs in field collected mosquitoes as revealed by a linkage disequilibrium analysis using the conservative analyses***

| | | | | | |
|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | -0.9274 | ** |
| **1** | **1** | **2** | **1** | **-0.2152** | |
| 1 | 2 | 1 | 3 | 1.6352 | ** |
| 1 | 2 | 2 | 0 | 0.1497 | |
| **1** | **3** | **1** | **6** | **-0.1412** | |
| 1 | 3 | 2 | 1 | -0.3766 | * |
| 2 | 1 | 1 | 7 | 0.5457 | ** |
| 2 | 1 | 2 | 0 | 0.1775 | |
| 2 | 2 | 1 | 7 | -1.1799 | ** |
| 2 | 2 | 2 | 0 | 0.1137 | |
| **2** | **3** | **1** | **4** | **0.1549** | |
| 2 | 3 | 2 | 0 | 0.1072 | |
| 3 | 1 | 1 | 0 | 0.9777 | ** |
| 3 | 1 | 2 | 0 | 0.0465 | |
| 3 | 2 | 1 | 2 | -1.7804 | ** |
| 3 | 2 | 2 | 0 | 0.0298 | |
| 3 | 3 | 1 | 0 | 0.5907 | ** |
| 3 | 3 | 2 | 0 | 0.0281 | |
| | | | 4 | | |
| $X^2_{lrt}$ Test | 12.6095 | | | | |
| p-value | 0.0133 | 4 d.f. | | | |

*The 18 possible haplotype combinations were analyzed for linkage disequilibrium with three S segments, three M segments and two L segments. The observed column is the number of that specific haplotype combination found in the population. The $R_{ijk}$ is the correlation coefficient for linkage disequilibrium. The highlighted samples are those haplotype combinations that are in linkage equilibrium that were observed in the population.

* p-value $\leq 0.05$, ** p-value $\leq 0.01$, *** p-value $\leq 0.001$

**Table 2.6: The S segment is the segment likely to reassort revealed by the linkage disequilibrium analysis of haplotype combinations of S*M, M*L, and S*L***

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 7 | -0.13306 | |
| 1 | 2 | 1 | 0.26647 | |
| 1 | 3 | 4 | -0.12059 | |
| 2 | 1 | 7 | -0.13306 | |
| 2 | 2 | 2 | 0.14922 | |
| 2 | 3 | 3 | 0.00000 | |
| 3 | 1 | 7 | 0.15066 | |
| 3 | 2 | 7 | -0.22207 | |
| 3 | 3 | 4 | 0.05462 | |
| 4 | 1 | 0 | 0.21336 | |
| 4 | 2 | 2 | -0.36464 | ** |
| 4 | 3 | 0 | 0.12892 | |
| | | 44 | | |
| | | | | |
| Chi-Squ | | | 9.54630 | 6 d.f. |
| p-value | | | 0.14510 | |

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 7 | -0.46579 | *** |
| 1 | 2 | 5 | 0.02688 | |
| 1 | 3 | 8 | 0.32302 | ** |
| 1 | 4 | 1 | -0.01016 | |
| 2 | 1 | 0 | 0.27255 | |
| 2 | 2 | 3 | 0.00000 | |
| 2 | 3 | 9 | -0.25741 | * |
| 2 | 4 | 0 | 0.13674 | |
| 3 | 1 | 0 | 0.25696 | |
| 3 | 2 | 3 | -0.03101 | |
| 3 | 3 | 7 | -0.10786 | |
| 3 | 4 | 1 | -0.12892 | |
| | | 44 | | |
| | | | | |
| Chi-Squ | | | 10.87789 | 6 d.f. |
| p-value | | | 0.09222 | |

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 2 | -0.01298 | |
| 1 | 2 | 3 | 0.00000 | |
| 1 | 3 | 5 | 0.16207 | |
| 1 | 4 | 2 | -0.36464 | ** |
| 2 | 1 | 1 | 0.12979 | |
| 2 | 2 | 6 | -0.36178 | ** |
| 2 | 3 | 5 | 0.16207 | |
| 2 | 4 | 0 | 0.13674 | |
| 3 | 1 | 4 | -0.14695 | |
| 3 | 2 | 2 | 0.27309 | |
| 3 | 3 | 12 | -0.20726 | |
| 3 | 4 | 0 | 0.18579 | |
| 4 | 1 | 0 | 0.09712 | |
| 4 | 2 | 0 | 0.12892 | |
| 4 | 3 | 2 | -0.20384 | |
| 4 | 4 | 0 | 0.04873 | |
| | | 44 | | |
| | | | | |
| Chi-squ | | | 13.40079 | 9 d.f. |
| p-value | | | 0.14529 | |

*The possible haplotype combinations (12: S*M, 12: M*L, and 16: S*L) were analyzed for linkage disequilibrium with four S segments, three M segments and four L segments. The observed column is the number of that specific haplotype combination found in the population. The $R_{ijk}$ is the correlation coefficient for linkage disequilibrium.
* p-value $\leq 0.10$, ** p-value $\leq 0.05$, *** p-value $\leq 0.01$

**Table 2.7: Minimal linkage disequilibrium was observed in the 2x2 linkage disequilibrium analysis of haplotype combinations of S*M, M*L, and S*L using the conservative clades.***

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 14 | -0.23802 | |
| 1 | 2 | 3 | 0.37181 | ** |
| 1 | 3 | 7 | -0.10786 | |
| 2 | 1 | 7 | 0.15066 | |
| 2 | 2 | 7 | -0.22207 | |
| 2 | 3 | 4 | 0.05462 | |
| 3 | 1 | 0 | 0.21336 | |
| 3 | 2 | 2 | -0.36464 | ** |
| 3 | 3 | 0 | 0.12892 | |
| | | 44 | | |
| | | | | |
| Chi-Squ | | | 9.22685 | 4 d.f. |
| p-value | | | 0.05567 | |

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 20 | 0.01016 | |
| 1 | 2 | 1 | -0.01016 | |
| 2 | 1 | 12 | -0.13674 | |
| 2 | 2 | 0 | 0.13674 | |
| 3 | 1 | 10 | 0.12892 | |
| 3 | 2 | 1 | -0.12892 | |
| | | 44 | | |
| | | | | |
| Chi-Squ | | | 1.09751 | 2 d.f. |
| p-value | | | 0.57767 | |

| | | | | |
|---|---|---|---|---|
| 1 | 1 | 22 | 0.20384 | |
| 1 | 2 | 2 | -0.20384 | |
| 2 | 1 | 18 | -0.18579 | |
| 2 | 2 | 0 | 0.18579 | |
| 3 | 1 | 2 | -0.04873 | |
| 3 | 2 | 0 | 0.04873 | |
| | | 44 | | |
| | | | | |
| Chi-Squ | | | 1.74603 | 2 d.f. |
| p-value | | | 0.41769 | |

*The possible haplotype combinations (9: S*M, 6: M*L, and 6: S*L) were analyzed for linkage disequilibrium with three S segments, three M segments and two L segments. The observed column is the number of that specific haplotype combination found in the population. The $R_{ijk}$ is the correlation coefficient for linkage disequilibrium.
* p-value $\leq 0.10$, ** p-value $\leq 0.05$, *** p-value $\leq 0.01$

support the evidence for reassortment, because none of the comparisons produced statistically significant results (Table 2.6 and 2.7).

**6. Occurrence of quasispecies**

There was also evidence of the existence of quasispecies in some of the mosquitoes analyzed. In these cases, alignment of sequences of three to five clones amplified from a single mosquito showed two or three different viral haplotypes for each segment (Figure 2.8). This occurred in 10/44 samples in the S segment, 25/44 samples in the M segment, and 18/44 samples in the L segment. In the remaining samples, there was a strong consensus sequence for a given segment with no variable sites. There was no variability in sequences from any of the three segments of the four virus isolates.

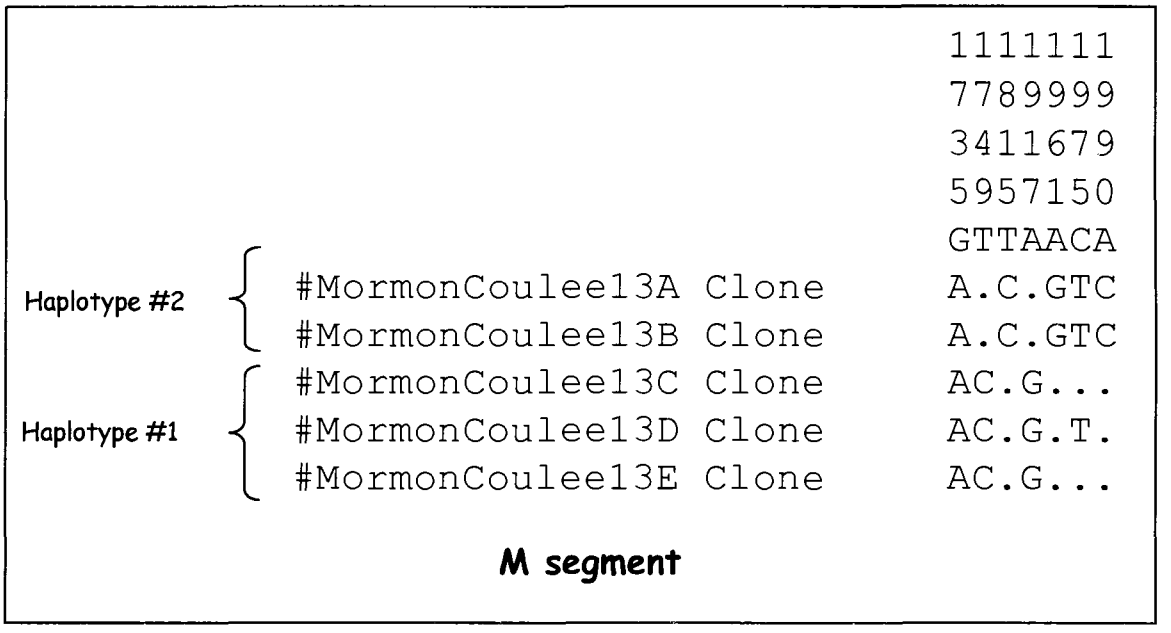



**Figure 2.8: Evidence for multiple virus haplotypes in a single mosquito**
**The viral sequence amplified from the mosquito collected from Mormon Coulee yielded both haplotypes 1 and 2 of the M segment.**

## D. DISCUSSION

Segment reassortment occurred in 25-38.7% of mosquitoes as determined by both phylogenetic analyses and linkage disequilibrium analysis (Figure 2.5-2.7 and Tables 2.4 and 2.5). The analyses for genome sequence polymorphism, nucleotide substitution, distance matrix, and sequence identity matrix revealed different haplotypes present in LACV-infected mosquitoes. The distribution of the sequences in the phylogenetic trees would be identical for all three segments if there were no reassortment. However, the phylogenetic trees are highly variable when comparing the S, M and L tree topologies (Figure 2.5-2.7). This analysis was performed without the use of a prototype LACV that is known to not have a reassorted genome, so therefore the samples that are or are not reassortants can not be determined through this analysis.

The two by two analyses of the possible haplotype combinations of S and M, M and L, and S and L suggests that the S segment could be the segment more likely to cluster independently than the M and L segments (Tables 2.6 and 2.7). The M and L analysis revealed linkage disequilibrium in 15 of the 44 samples (34.1%) compared to 4.5% (2/44) of the S and M haplotype combinations and 18.1% (8/44) of the S and L haplotype combinations. The $\chi^2_{Ind}$ test further supports this at a 90% confidence interval. Interestingly, the reassortant virus isolated by Klimas *et al.* contained an M and L segment of one genotype of LACV and an S segment of a different genotype (Klimas et al., 1981). One possible reason for this observation could be due to the virulence determinant properties of both the M and L segments. The M segment codes for the glycoproteins, G1 and G2, and conditions attachment and entry into cells, tropisms, and mosquito infectivity (Gonzalez-Scarano et al., 1992). The L segment codes for the RNA-

dependent-RNA polymerase and conditions replication efficiency (Rust et al., 1999). Therefore these two segments may need to be in linkage disequilibrium to ensure infection of a host.

There was evidence for the presence of quasispecies in LACV RNA isolated from mosquitoes (Figure 2.8). Quasispecies occur when multiple haplotypes are available for packaging. There may be a haplotype that has a greater fitness in one environment compared to another and this mechanism could allow LACV to have a successful transmission cycle. Quasispecies might have developed to allow LACV to adapt to the different landscapes in the host and the vector. Even though a mosquito may contain two or three haplotypes of the LACV RNA, only certain haplotypes may be packaged, resulting in infectious particles. The haplotype packaged depends on the environment and the replicative fitness. The presence of quasispecies in this analysis therefore demonstrates the evolutionary potential of LACV.

Although the LACV RNA isolated from mosquitoes provides evidence for reassortment, the occurrence of quasispecies and natural sequence variation confounds this interpretation. The amount of reassortment detected in this study could possibly be the effect of analyzing pre-selection phase quasispecies RNA. In order to determine if infectious virus has genome segments demonstrating quasispecies-like sequence variability, viruses will need to be isolated from mosquitoes, preferably from saliva or ovaries, and plaque-purified clones will need to be sequenced.

Results of the experiments described here suggest that segment reassortment in LACV occurs frequently in the field. One possible reason for this could be the ability of *Ae. triseriatus* mosquitoes to become dually infected; 100% of mosquitoes ingesting two

different LACV isolates simultaneously or within four hours become dually infected (Beaty et al., 1985). Even at 48 hours post-initial bloodmeal, 27% of mosquitoes ingesting a second virus still become dually infected before a barrier to superinfection develops. In addition, when TO-infected mosquitoes ingested a bloodmeal containing a heterologous LACV, 18.6% became dually infected (Borucki et al., 1999). These experiments suggest that dual infection can frequently occur through both oral and transovarial infection and therefore increase the possibility of segment reassortment of LACV in vectors. The newly evolved viruses are also efficiently transmitted (Beaty et al., 1985). These experiments were performed in a controlled laboratory setting, but they demonstrate the potential for segment reassortment to occur frequently in nature.

Because LACV is already the leading cause of pediatric encephalitis in the United States, there are important epidemiologic implications of reassortment in LACV-infected *Ae. triseriatus* mosquitoes in the field. If two viruses reassort and create a new virus, vital determinants of the pathogenesis and transmission cycle could be altered. New viral phenotypes could be capable of infecting new vector species or vertebrate hosts. New viruses could also be introduced into new arbovirus cycles with potentially significant epidemiological consequences (Beaty & Calisher, 1991). For example, the geographic distribution of LACV is currently determined by the distribution of *Ae. triseriatus* mosquitoes and chipmunks and tree squirrels. If a new viral phenotype infected birds, the geographic distribution could easily be increased by migratory patterns, similar to the observed geographic spread of West Nile virus (Family: *Flaviviridae*, Genus: *Flavivirus*) in the United States (Gubler, 2002). If a new virus established a new transmission cycle with a different mosquito that more aggressively feeds on humans, increased human

infection could occur, possibly becoming clinically significant in both adults and children. In addition, a new viral phenotype could be more virulent and exhibit different tropisms, perhaps causing humans to develop a high enough viremia titer to infect biting mosquitoes and/or to cross the blood-brain barrier more efficiently. The implications of reassortment in nature are poorly understood. Determination of the evolutionary potential of LACV through genetic shift may permit prediction of the epidemiologic consequences of these events.

# III. GENOTYPIC VARIABILITY OF LA CROSSE VIRUS: INVESTIGATIONS OF GEOGRAPHICAL (SPATIAL), ENVIRONMENTAL, AND TEMPORAL FACTORS THAT CONDITION THE GENETIC STRUCTURE OF THE VIRUS IN AN ENDEMIC AREA

## A. INTRODUCTION

La Crosse virus (LACV) belongs to the California serogroup of the genus *Orthobunyavirus* in the family *Bunyaviridae*. It is the primary cause of pediatric arboviral encephalitis in the United States. The incidence of LACV is 5-10 cases per 100,000 in endemic areas. LACV encephalitis afflicts mainly children between the ages of 3 and 15 with the majority of the cases reported in the Midwestern United States (Rust et al., 1999). However, LACV has been isolated from 13 states in the eastern U.S. extending from Texas to Minnesota and from New York to Georgia (Calisher, 1994). Clinical cases of LACV have been reported in 28 states (CDC, 2005, Grayson & Calisher, 1983).

The genotypic variation of LACV isolated from different geographic regions of the United States and different ecological niches has been examined through oligonucleotide fingerprints (El Said et al., 1979, Klimas et al., 1981). Viruses were isolated from the Midwestern (Wisconsin, Minnesota, Illinois, Indiana, and Ohio), Eastern (New York, Georgia and North Carolina) and Southern (Texas) portions of the United States. Comparison of oligonucleotide fingerprints revealed three haplotypes of the virus. Type A LACV was recovered from Wisconsin, Minnesota, Indiana and Ohio. Type B LACV was isolated from Minnesota, Wisconsin and Illinois. The A and B type viruses are sympatric in the upper Midwest. Type C was isolated from the eastern and southern regions of the United States. An additional study compared LACV isolated from a variety of mosquito species in the northern United States to the prototype LACV isolated from the first fatal case in 1960 (El Said et al., 1979, Thompson et al., 1965). Both studies found that no two viruses yielded identical fingerprints, suggesting there has been considerable evolution of LACV in nature. There was variation between the

different isolates within the type A and B haplotypes and within viruses from the different ecological niches. This probably resulted from genetic drift, which is a major contributor to the evolutionary potential of *Bunyaviridae* viruses (Beaty & Calisher, 1991, Bishop & Beaty, 1988).

LACV isolates from mosquitoes and humans have also been analyzed using more sensitive techniques such as sequencing and phylogenetic analysis to investigate genetic variation over time and distance (Huang et al., 1995, Huang et al., 1997). The M segment of LACV isolates from the brains of two children autopsied 18 years apart in Wisconsin were compared (Huang et al., 1995). There was an overall identity of 99.6% in the nucleotide sequence. An additional phylogenetic study included a LACV isolate recovered from a fatal case of encephalitis that occurred in Missouri in 1993 and multiple mosquito LACV isolates (Huang et al., 1997). This allowed an opportunity to compare sequence data from viruses that were separated spatially (upper Midwest and Missouri) and chronologically (33 years). Comparison of M segment genomic RNA from these viruses showed that the nucleotide and deduced amino acid sequences of the human isolates were highly conserved. The high degree of conservation over time and space led to the hypothesis that human infections of LACV are most likely related to a specific genotype that resulted in severe disease and a fatal outcome. Alternatively, LACV passage in and isolation from the central nervous system of humans could select for specific virus genotypes. Clearly, the hypothesis of fatal human infections being associated with a narrow range of genotypes needs to be investigated more rigorously. It is noteworthy that the M segment sequence isolated from mosquitoes was similar to that

of the viruses recovered from humans, suggesting that the genotype associated with human fatalities also circulates in mosquitoes (Huang et al., 1997).

An additional phylogenetic analysis of the LACV M segment demonstrated three different lineages of the virus (Armstrong & Andreadis, 2006). One lineage consisted of viruses isolated from the Midwest (Wisconsin, Minnesota, Missouri, and Ohio) and Appalachian regions (Tennessee, North Carolina, and West Virginia). The substructure of this lineage revealed four homogeneous clusters of variants from (1) Minnesota, western Wisconsin and Missouri, (2) Wisconsin and western Ohio, (3) West Virginia, and (4) North Carolina and Tennessee. A second lineage consisted of viruses isolated from southeastern (Alabama and Georgia) and northeastern (New York) United States. A third lineage consisted solely of a new genetic variant of LACV isolated from Connecticut. These studies support the previous observations (Klimas et al., 1981) that the viruses loosely cluster into different lineages based upon a geographical distance. Overall, these analyses confirmed the evolutionary diversification among LACV in nature. The epidemiological significance of this genetic variability remains to be determined.

The La Crosse, Wisconsin, region, including nearby areas in southeast Minnesota and northeast Iowa, is arguably one of the most intensely studied regions in terms of LACV molecular epidemiology, beginning with the studies by Klimas *et al.* (Klimas et al., 1981) covered in I. (F.2.A, F.2.B.). However much remains to be determined concerning the genetic structure of viruses in the area. The geographical, ecological and temporal determinants of genetic structure and gene flow in the virus population needs to be investigated. Previous studies have demonstrated that there are no barriers to gene flow in *Ae. triseriatus* mosquitoes in the collection area in the La Crosse, Wisconsin

region, southeast Minnesota and northeast Iowa (Beck et al., 2005). Analysis of the mitochondrial gene, NADH dehydrogenase subunit 4, found no evidence for genetic isolation by distance and Interstate 90 and the Mississippi River were not barriers to gene flow (Beck et al., 2005). Since *Ae. triseriatus* mosquitoes exist as a panmictic population it was important to determine if there were barriers to LACV viral gene flow in the same collection area.

Many reports have demonstrated that viruses from the *Bunyaviridae* family are evolving geographically, ecologically and chronologically (Armstrong & Andreadis, 2006, Avsic-Zupanc et al., 2007, Bird et al., 2007, Deyde et al., 2006, El Said et al., 1979, Huang et al., 1996, Klimas et al., 1981, Vapalahti et al., 1996, Yamakawa et al., 2006). Clearly, there is regional genetic isolation of LACV as determined by both oligonucleotide fingerprinting and more recently by sequence analysis (Armstrong & Andreadis, 2006, El Said et al., 1979, Huang et al., 1995, Huang et al., 1997, Klimas et al., 1981). However, the genetic variability of LACV on a smaller spatial scale needs to be investigated.

In this study, the genetic variation of LACV was investigated in terms of geographic origin, environmental terrain of the collection site, and collection year in a study range of 15,360 $km^2$ in western Wisconsin, southeastern Minnesota and northeastern Iowa. The principal goal of this study was to characterize the geographic (spatial) variability and viral gene flow of LACV in an endemic region using sensitive molecular epidemiological techniques. A phylogenetic analysis of LACV RNA isolated from field-infected *Ae. triseriatus* mosquitoes collected from Wisconsin and Minnesota was executed. The hypothesis for this study was that genetic variation of LACV is

associated with the geographic origin, and distinct geographic lineages will be discovered on a smaller spatial scale than previously studied. Possible physical barriers to viral gene flow that could prevent the exchange of viruses in the collection region were also investigated. The outcome of this study could have epidemiological implications. As previously mentioned, there may be a narrow range of genotypes correlated with severe disease and fatal outcomes (Huang et al., 1997). The presence and potential trafficking of such genotypes could pose major risks to humans and detection of such genotypes could lead to targeted control efforts.

## B. MATERIALS AND METHODS

### 1. Egg collection

*Aedes triseriatus* eggs were collected from five oviposition traps in each of 151 sites in Minnesota (n = 37), Wisconsin (n = 108) and Iowa (n = 6) in areas where LACV encephalitis cases occurred or areas that contained clusters of people judged by the La Crosse County Health Department to be at risk for LACV infection (e.g. wooded areas adjacent to houses with children, schools, or playgrounds). Each trap consisted of a can (6.5 cm x 11 cm) painted black, half filled with tap water, with a seed germination paper lining the inside perimeter, placed slightly above ground level (Beck et al., 2005). Egg papers were collected from traps after 10 days and sent to the insectaries at the Arthropod-borne and Infectious Diseases Laboratory (AIDL) at Colorado State University (CSU), Fort Collins, CO. Mosquito eggs that had entered diapause in the fall of 2000 were collected in the spring of 2001. Mosquito eggs were also collected between mid-June and August of 2004 and 2005. Eggs were collected in Crawford, La Crosse, Monroe, Vernon, Lafayette and Iowa counties in Wisconsin, Winona and Houston

counties in Minnesota and Clayton County in Iowa (Figure 3.1). The eggs collected in 2001, 2004 and 2005 were hatched immediately upon delivery and reared to adults.

**2. Immunofluorescence assay**

To determine infection status, mosquito heads were severed, squashed onto acid-washed microscope slides, and fixed in acetone. They were assayed for LACV antigen by direct immunofluorescence assay (IFA) using a LACV-specific polyclonal antibody (Beaty & Thompson, 1975).

**3. RNA purification amplification by reverse transcription-PCR**

The posterior half of each mosquito abdomen was individually homogenized in 500µl of Trizol (Invitrogen), using a pellet pestle (Fisher Scientific), then total RNA was extracted according to manufacturer's instructions. Portions of the LACV S and M RNA segments were transcribed to cDNA using Superscript II reverse transcriptase (Invitrogen) and amplified by PCR using Ex Taq DNA polymerase (Takara) according to manufacturer's instructions. The primers specific for the S segment (forward: 5'-GCAA ATGGATTTGATCCTGATGCAG-3', reverse: 5'-CTTAAGGCCTTCTTCAGGTAT TGAG-3') amplified a 461 nucleotide region (nucleotides 144 to 605) of the nucleocapsid gene that was selected because it was the most variable region of the published S segment sequences. The S segment is 984 nucleotides in length, so the amplified region encompasses almost half the entire segment. The primers specific for the M segment (forward: 5'-CCAAAAGCAACA AAAGAAAGA-3', reverse: 5'- CTGA AGGCATGATGCAAAG-3') amplified a highly variable 410 nucleotide region (nucleotides 1585 to 1995), in the 5' half of the G1 gene (Borucki et al., 2001).

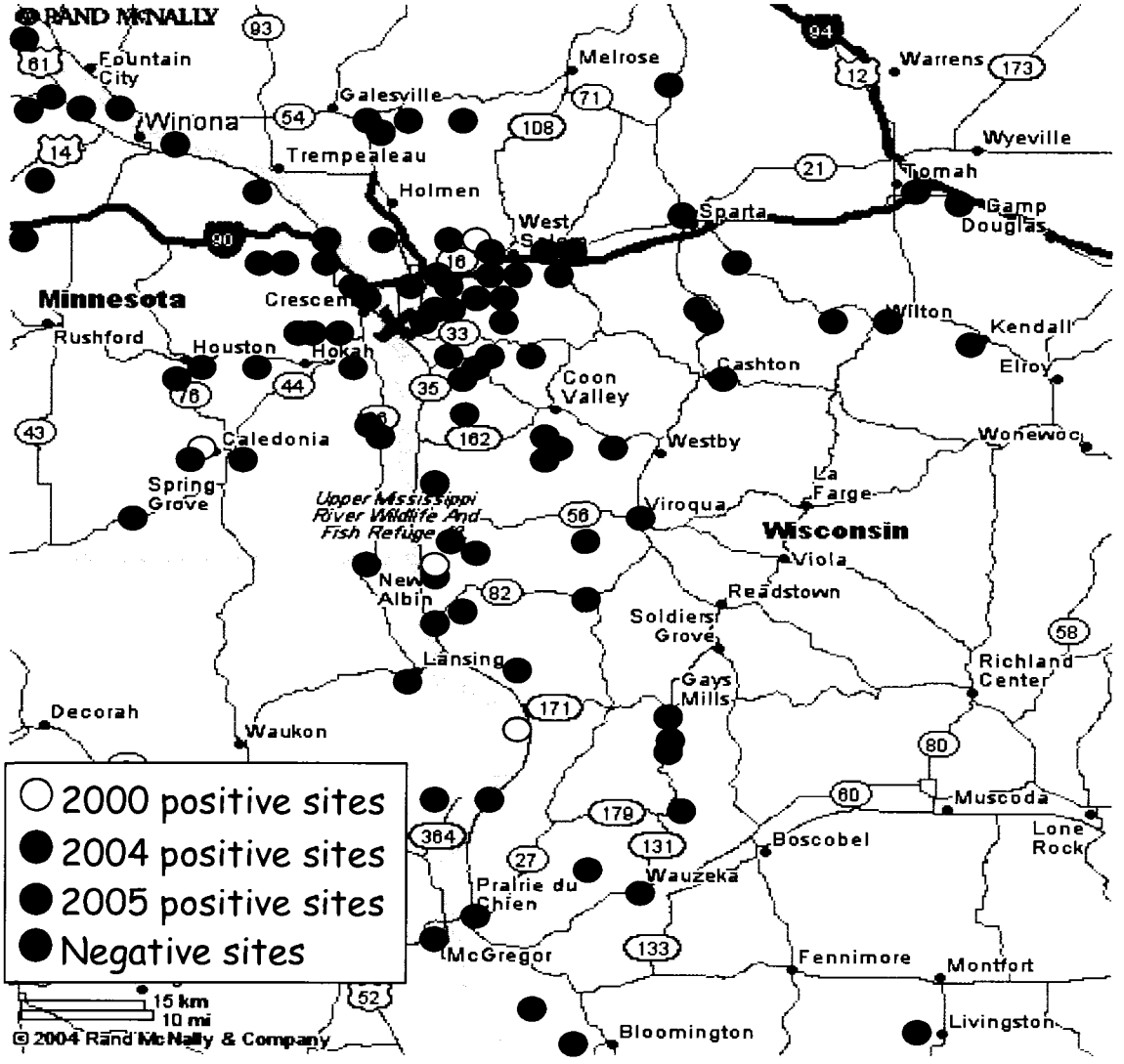



**Figure 3.1: Mosquito collection sites in Minnesota, Wisconsin and Iowa.**
Circles represent all collection sites. Yellow circles are the sites where LACV positive mosquitoes were collected in 2000, red circles are the sites where LACV positive mosquitoes were collected in 2004, blue circles are the sites where LACV positive mosquitoes were collected in 2005, and black circles were the sites without LACV positive mosquitoes.

Primers specific for the *Ae. triseriatus* ribosomal protein Rpl34 mRNA were used to amplify the positive control. PCR was performed as follows: 94°C for 5 minutes, 35 cycles of [94°C for 1 minute, 55°C for 1 minute and 72°C for 1 minute] followed by a final extension at 72°C for 8 minutes.

**4. Amplicon cloning and sequencing**

PCR products were separated electrophoretically in 1% agarose gels with TAE (Tris-Acetate-EDTA) buffer, stained with ethidium bromide, excised and extracted using a kit from Marligen Biosciences according to manufacturer's instructions. PCR products were inserted into the pCR4-TOPO cloning vector (Invitrogen) and resulting plasmids were used to transform competent TOP10 *E. coli* cells (Invitrogen). Cells were grown on LB agar containing ampicillin (50μg/ml) and kanamycin (50μg/ml). Colonies were screened for inserts by PCR amplification, using the original primers and positive products were purified using a QIAquick spin column (Qiagen). Three to five cDNA clones per segment from each mosquito were sequenced using the ABI PRISM dye terminator cycle sequencing kit (Applied Biosystems, CA, USA) and the ABI 310 DNA automated sequencer at Macromolecular Resources, CSU. A 415 nucleotide region of the S segment (nucleotides 190-604) and a 358 nucleotide region of the M segment (nucleotides 1637-1994) were sequenced and analyzed.

**5. AMOVA analysis**

The genetic structure of a population is investigated by an analysis of variance framework (Cockerham, 1969, Cockerham, 1973, Long, 1986, Weir & Cockerham, 1984). The Analysis of Molecular Variance (AMOVA) approach used in Arlequin version 2.0 (University of Geneva, Geneva, Switzerland) is based on analyses of variance

of gene frequencies and accounts for the number of mutations between molecular haplotypes (Excoffier et al., 1992). Genetic structure is tested by defining groups of populations. A hierarchical analysis of variance partitions the total variance into covariance components due to intra-individual differences, inter-individual differences, and/or inter-population differences.

Three fixation indices are calculated to determine the significance of the observed variation:

(1) $F_{\text{among mosquitoes in a collection}}$

(2) $F_{\text{among collections in each group}}$

(3) $F_{\text{among the groups}}$

$F_{\text{among mosquitoes in a collection}}$ is a measure of the sequence variation observed within a collection site compared to the sequence variation observed in the entire population. $F_{\text{among collections in each county}}$ measures the sequence variation observed between the collection sites within a group compared to the sequence variation observed in the entire population. $F_{\text{among the counties}}$ measures the sequence variation observed between groups (Hartl & Clark, 1997). The fixation indices range from 0 (indication of no difference between the overall population and its subpopulations) to 1 (indicating a completely heterogeneous subpopulation). The higher the value of F, the more significant variation observed. An AMOVA analysis was performed by grouping collection sites for four different associations, (1) by collection site county, (2) by quadrants as defined in Beck *et al.* (Beck et al., 2005), (3) by environment, and (4) by collection year.

**6. Determination of land cover types**

Google Earth (http://earth.google.com/) was used to visualize and classify the mosquito sampling locations into four different general land covers. The specific latitude and longitude of each sampling point were entered into Google Earth and associated satellite images were viewed. A qualitative description of the land cover of each collection site was used to classify the sites into four distinct classes: (1) Patchy landscape (composed of agriculture) within 0.5 km distance of a small water source, (2) mixed forest within a 1 km distance of the Mississippi River, (3) residential/urban area, and (4) mixed forest not proximal to water sources.

**7. Distance Matrix and Phylogenetic Analysis**

A distance matrix was created using the Tamura-Nei distance. A genetic distance is measured among all nucleotides and results in associated assumptions of relatedness. Tamura-Nei considers the probability that any nucleotide will change to one specific nucleotide, subdivides transitions into purines and pyrimidines and considers the frequency of each nucleotide separately (Black & Salman, 2005). Maximum parsimony phylogenetic analysis was performed using the Phylogenetic Analysis Using Parsimony (PAUP) 4.0b10 package (Swofford, 2003). The phylogenetic trees indicate the branches that appear in the majority of the 100 bootstrap replications and the frequency with which these appear among the replications.

**8. Haplotype distribution**

Sequence ID matrix, Tamura-Nei distance matrix, nucleotide substitution and maximum parsimony phylogenetic analysis were used to separate the sequences into different haplotypes for each segment.

## C. RESULTS

### 1. Study population

Eighty-seven LACV-RNA positive field-infected mosquitoes were analyzed – 66 females and 21 males. There were 13 sites with more than 1 positive mosquito and 10 collection sites yielded positive mosquitoes in multiple years. Viral RNA was extracted from field-infected mosquitoes and the S and M segments were amplified by RT-PCR and analyzed through sequencing. In total 81 S segment and 69 M segment sequences were analyzed (Table 3.1).

### 2. Haplotype determination

The S segment clustered into 5 haplotypes based on 5 nonsynonymous mutations and the M segment clustered into 6 haplotypes based on 6 nonsynonymous mutations (Figure 3.2).

### 3. AMOVA for determination of isolation by distance

The data were analyzed by grouping collection sites by the counties in which they were located. The LACV haplotype frequencies were compared within collection sites, between collection sites in a county, and between the counties. This analysis was performed to determine if there is isolation by distance observed for the S and M segments in the 15,360 $km^2$ study range.

#### a. S segment

76.25% of the variation was seen within the collection sites ($p < 0.01$) while 17.40% of the variation was seen between collections within the counties ($p > 0.05$) and very little variation (6.35%) was seen between the counties ($p > 0.05$). These results indicate that there was no evidence of isolation by distance with the S segment of LACV

**Table 3.1. LACV-infected field collected mosquitoes from Wisconsin, Minnesota, and Iowa collected in 2000, 2004 and 2005***

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| PP/Clayton/2005 | Clayton | IA | 2005 | 2 | SW | Flat land/Near Miss river | x | x | |
| GAY120/Crawford/2005 | Crawford | WI | 2005 | 1 | SE | Hills/streams | x | | |
| GAY120/Crawford/2004 | Crawford | WI | 2004 | 1 | SE | Hills/streams | x | x | x |
| GAY724/Crawford/2005 | Crawford | WI | 2005 | 1 | SE | Hills/streams | x | x | |
| HWYBBS/Crawford/2005 | Crawford | WI | 2005 | 1 | SE | Flat land/Near Miss river | x | x | |
| LRHE/Crawford/2005 | Crawford | WI | 2000 | 1 | NE | Flat land/Near Miss river | x | x | x |
| NAT/Crawford/2004 | Crawford | WI | 2004 | 1 | NE | Hills | x | x | x |
| PDC/Crawford/2004 | Crawford | WI | 2004 | 1 | NE | Flat land/near city | | | x |
| RCS/Crawford/2004 | Crawford | WI | 2004 | 1 | NE | Hills/streams | x | x | x |
| RCS/Crawford/2004 | Crawford | WI | 2004 | 4 | NE | Hills/streams | x | | |
| WKCS/Crawford/2004 | Crawford | WI | 2004 | 2 | NE | Hills/streams | x | | |
| WKCS/Crawford/2004 | Crawford | WI | 2004 | 1 | NE | Hills/streams | x | x | x |
| PLT/Grant/2005 | Grant | WI | 2005 | 5 | SE | Flat land/near city | x | x | |
| BRS/Houston/2005 | Houston | MN | 2005 | 1 | SW | Flat land/Near Miss river | x | x | |
| BRS/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flat land/Near Miss river | x | x | x |
| BWV/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flat land/Near Miss river | x | x | x |
| CAL-B/Houston/2000 | Houston | MN | 2000 | 2 | SE | Flat land | x | x | x |
| CAL-B/Houston/2000 | Houston | MN | 2000 | 1 | SE | Flat land | x | x | |
| CAL-B/Houston/2000 | Houston | MN | 2000 | 3 | SE | Flat land | | | x |
| Cal-B/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flat land | x | x | x |
| CAL-D/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flat land | x | x | x |
| CAL-GA/Houston/2005 | Houston | MN | 2005 | 1 | SW | Flat land | x | x | |
| CAL-GA/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flat land | x | x | x |
| HC/Houston/2004 | Houston | MN | 2004 | 1 | SW | Hills | x | | |
| HCS/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flatland | x | x | x |
| HHS/Houston/2005 | Houston | MN | 2005 | 1 | SW | Flat land | x | x | |
| HHS/Houston/2004 | Houston | MN | 2004 | 1 | SW | Flatland | x | x | x |
| GRL/La Crosse/2004 | La Crosse | WI | 2004 | 1 | NE | Hills | x | x | x |
| INNSB/La Crosse/2000 | La Crosse | WI | 2000 | 1 | SE | Hills | x | x | x |
| INNSL/La Crosse/2004 | La Crosse | WI | 2004 | 1 | SE | Hills | x | x | x |
| INNSL/La Crosse/2005 | La Crosse | WI | 2005 | 1 | NE | Hills | x | x | |
| LAXCC/La Crosse/2005 | La Crosse | WI | 2005 | 1 | SE | Hills | x | x | |
| LAXCC/La Crosse/2004 | La Crosse | WI | 2004 | 1 | SE | Hills | x | x | x |
| MCBB/La Crosse/2004 | La Crosse | WI | 2004 | 1 | NE | Hills | x | x | x |
| MCP/La Crosse/2004 | La Crosse | WI | 2004 | 2 | NE | Hills | x | x | x |
| OTS/La Crosse/2004 | La Crosse | WI | 2004 | 1 | SE | Flat land/near miss river | x | x | x |
| SRS/La Crosse/2004 | La Crosse | WI | 2004 | 1 | NE | Flats/near city | x | x | x |
| SST/La Crosse/2004 | La Crosse | WI | 2004 | 1 | SE | Hills/streams | x | x | x |
| WS4/La Croose/2005 | La Crosse | WI | 2005 | 1 | NE | Hills | x | x | |
| WSB/La Crosse/2004 | La Crosse | WI | 2004 | 1 | SE | Flats/near city | x | x | x |
| BEN/Layfayette/2005 | Layfayette | WI | 2005 | 5 | SE | Flat land/near city | x | x | |
| CAT/Monroe/2004 | Monroe | WI | 2004 | 1 | SE | Hills | x | x | x |
| MEL/Monroe/2005 | Monroe | WI | 2005 | 2 | NE | Hills/streams | x | | |
| OAK/Monroe/2005 | Monroe | WI | 2005 | 2 | SE | Flat land | x | x | |
| TFP/Monroe/2004 | Monroe | WI | 2004 | 1 | SE | Flat land/near city | x | x | x |
| WBF/Monroe/2004 | Monroe | WI | 2004 | 1 | NE | Hills | x | x | x |
| ESO/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Hills/streams | x | x | x |
| ESO/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Hills/streams | x | | |
| GSO/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Wooded/Mississippi | x | | |
| H0/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Hills/streams | x | x | x |
| MVD/Vernon/2005 | Vernon | WI | 2005 | 1 | NE | Hills | x | x | |
| SHR/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Hills/streams | x | x | x |
| SVP/Vernon/2005 | Vernon | WI | 2005 | 1 | NE | Wooded/Mississippi | x | x | |
| SVP/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Wooded/Mississippi | x | x | x |
| SVP/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Wooded/Mississippi | x | x | |
| VS/Vernon/2004 | Vernon | WI | 2004 | 1 | NE | Wooded/Mississippi | x | x | x |
| VS/Vernon/2000 | Vernon | WI | 2000 | 1 | SE | Wooded/Mississippi | x | x | x |
| VS/Vernon/2000 | Vernon | WI | 2000 | 1 | SE | Wooded/Mississippi | | | x |
| BC/Winona/2004 | Winona | MN | 2004 | 2 | NW | Wooded/Mississippi | x | x | x |
| DAK-90/Winona/2004 | Winona | MN | 2004 | 1 | NW | Wooded/Mississippi | x | x | x |
| HID/Winona/2005 | Winona | MN | 2005 | 2 | NW | Flats/near city | x | x | |
| LHS/Winona/2005 | Winona | MN | 2005 | 1 | SW | Flats/near city | x | x | |
| NFCS/Winona/2005 | Winona | MN | 2005 | 2 | SW | Flat land/near city | x | x | |
| NFCS/Winona/2004 | Winona | MN | 2004 | 1 | SW | Flat land/near city | x | x | x |
| RRA/Winona/2004 | Winona | MN | 2004 | 1 | SW | Flat land/Near Miss river | x | x | x |

*Different environmental terrains were defined as follows: (1) Patchy landscape within 0.5 km distance of a small water source, (2) mixed forest within a 1 km distance of the Mississippi River, (3) residential/urban area, and (4) mixed forest without a water source.

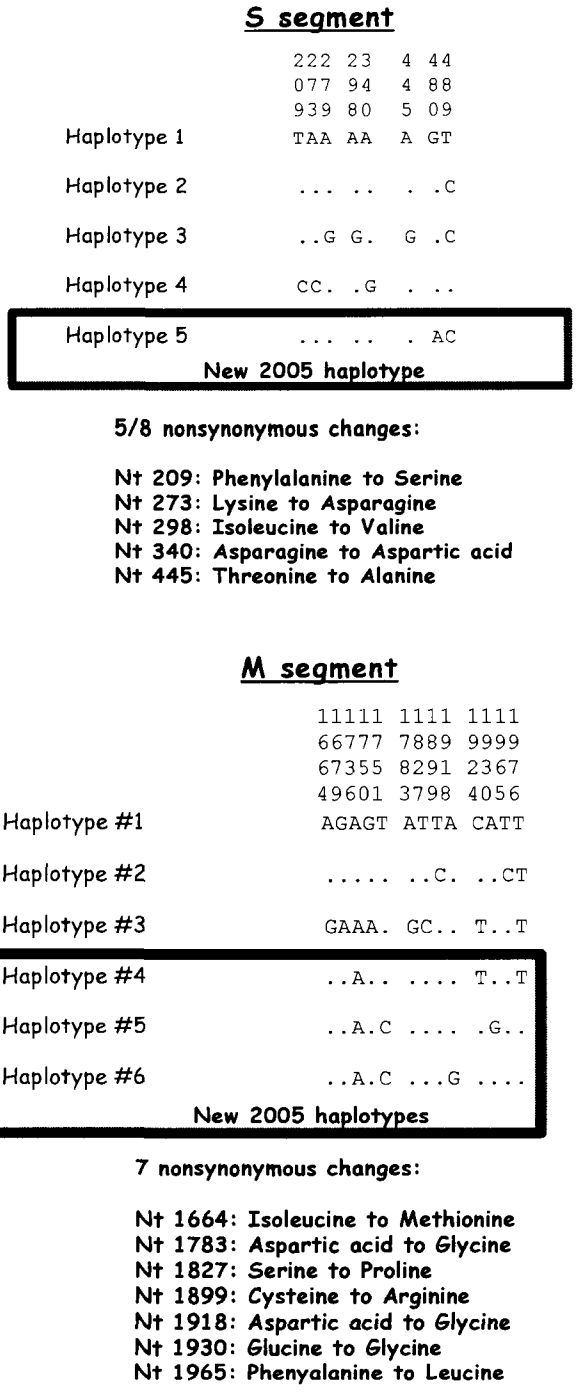



**Figure 3.2: Haplotype distribution for the S and M segment sequences of LACV amplified from mosquitoes collected in 2000, 2004, and 2005**
The segment sequence variation of the S segment with 5 haplotypes (A) and the M segment with 6 haplotypes (B). The genome position is provided above the genetic sequence.

RNA (Table 3.2) and all the S segment haplotypes can be expected to circulate throughout the 15,360 $km^2$ study range.

**b. M segment**

As with the S segment, most of the variation in virus sequences (73.63%) was seen within the collection sites ($p < 0.01$), while some sequence variation was detected among the collections within the counties (26.07%, $p < 0.05$) and very little variation between the counties (0.30%, $p > 0.05$) (Table 3.3). These results indicate that there was no evidence of isolation by distance with the M segment of LACV and all the M segment haplotypes can be expected to circulate throughout the 15,360 $km^2$ study range.

**4. AMOVA for determination of physical barriers to viral gene flow**

The data were analyzed by grouping by the quadrants in which the mosquitoes were collected; northeast, northwest, southeast and southwest (Table 3.1). This analysis was performed to determine if the Mississippi River and/or Interstate 90 are physical barriers to viral gene flow.

**a. S segment**

The majority of the sequence variation was found within the collection sites (67.53%, $p < 0.001$) and among the collection sites (29.18%, $p > 0.05$) as compared to very little sequence variation observed among the quadrants (3.30%, $p > 0.05$) (Table 3.4). These results indicate that I-90 and the Mississippi River are not barriers to viral gene flow in this area and that all LACV S segment haplotypes can be expected to circulate effectively through the study area.

**b. M segment**

Most of the observed sequence variation was detected within the collection sites (59.99%, $p < 0.001$) and between the collection sites (42.59%, $p < 0.001$). Little variation was observed between the quadrants (-2.57%, $p > 0.05$) (Table 3.5). These results indicate that I-90 and the Mississippi River are not barriers to viral gene flow in this area and that all LACV M segment haplotypes can be expected to circulate effectively through the study area.

**5. AMOVA for determination of ecological determinants**

This analysis was performed to determine if the environmental terrain of the collection site is a determinant of the genetic sequence. The haplotype frequencies were partitioned using AMOVA within the collections, among the collections within a specific environmental terrain and among environmental terrains. The four distinct classes of land cover were: (1) Patchy landscape (composed of agriculture) within 0.5 km distance of a small water source, (2) mixed forest within a 1 km distance of the Mississippi River, (3) residential/urban area, and (4) mixed forest not proximal to water sources (Table 3.1).

**a. S segment**

There was significant sequence variation observed within the collection sites (68.65%, $p < 0.001$). Some variation was observed between the different land cover groups (11.28%, $p < 0.01$) and among the collection sites within different land cover groups (19.77%, $p > 0.05$) (Table 3.6). These results suggest that the S segment haplotypes may be determined by the environmental terrain of the collection site.

**Table 3.2: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all nine counties**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 9 counties | 6 | 3.910 | 0.02443 | 6.35 |
| Among collections in each county | 32 | 12.933 | 0.06695 | 17.40 |
| Among mosquitoes in a collection | 28 | 8.217 | 0.29345 | **76.25**** |
| Total | 66 | 25.060 | 0.38483 | |

Fixation Indices: $F_{\text{among counties}} = 0.06348$
$F_{\text{among collections in each county}} = 0.18577$
$^{**}F_{\text{among mosquitoes in a collection}} = \mathbf{0.23746}$

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**Table 3.3: Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all nine counties**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 9 counties | 6 | 2.775 | 0.00113 | 0.30 |
| Among collections in each county | 29 | 12.139 | 0.09785 | **26.07*** |
| Among mosquitoes in a collection | 19 | 5.250 | 0.27632 | **73.63**** |
| Total | 54 | 20.164 | 0.37530 | |

Fixation Indices: $F_{\text{among counties}} = 0.00301$
$^{*}F_{\text{among collections in each county}} = \mathbf{0.26152}$
$^{**}F_{\text{among mosquitoes in a collection}} = \mathbf{0.26375}$

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**Table 3.4: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all four quadrants**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 4 quadrants | 3 | 2.158 | 0.01277 | 3.30 |
| Among collections in each quadrant | 43 | 19.642 | 0.11302 | 29.18 |
| Among mosquitoes in a collection | 36 | 9.417 | 0.26157 | **67.53*** ** |
| Total | 82 | 31.217 | 0.38736 | |

Fixation Indices: $F_{\text{among counties}}$ = 0.03297
$F_{\text{among collections in each county}}$ = 0.30171
***$F_{\text{among mosquitoes in a collection}}$ = **0.32473**

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**Table 3.5: Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all four quadrants**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 4 quadrants | 3 | 1.349 | -0.01030 | -2.57 |
| Among collections in each quadrant | 38 | 19.084 | 0.17037 | **42.59*** ** |
| Among mosquitoes in a collection | 25 | 6.000 | 0.24000 | **59.99*** ** |
| Total | 66 | 26.433 | 0.40008 | |

Fixation Indices: $F_{\text{among counties}}$ = -0.02574
***$F_{\text{among collections in each county}}$ = **0.41517**
***$F_{\text{among mosquitoes in a collection}}$ = **0.40012**

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**b. M segment**

Significant sequence variation was observed within the collection sites (59.31%, $p < 0.001$) and between the collection sites within different land cover groups (39.45%, $p < 0.001$). There was minimal sequence variation observed between land cover groups (1.24%, $p > 0.05$) (Table 3.7). These results suggest that the environmental terrain is not a determinant of LACV M segment haplotype.

**6. AMOVA for determination of temporal effects of LACV haplotypes**

This analysis was performed to determine if the year an infected mosquito was collected is a determinant of the LACV haplotype. The haplotype frequencies were partitioned using AMOVA within collections, between collections within a collection year, and between collection years.

**a. S segment**

There was significant sequence variation observed within the collection sites (36.75%, $p < 0.001$), between the collection sites within the same collection year (26.50%, $p < 0.001$) and between the years collected (36.75%, $p < 0.001$) (Table 3.8). These results suggest that the collection year may be a determinant of the S segment haplotype.

**b. M segment**

As with the S segment, all three of the observed sequence variations were significant. There was 31.63% of the variation observed within the collection sites ($p < 0.001$), 43.93% of the variation observed among the collection sites ($p < 0.001$), and

**Table 3.6: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all four environmental terrain categories**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 4 environments | 5 | 5.245 | 0.04384 | **11.28**** |
| Among collections in each environment | 41 | 16.362 | 0.07683 | 19.77 |
| Among mosquitoes in a collection | 37 | 9.917 | 0.26802 | **68.95***** |
| Total | 83 | 31.524 | 0.38869 | |

Fixation Indices: **$F_{\text{among land cover}}$ = **0.11279**
$F_{\text{among collections in each land cover}}$ = 0.22280
***$F_{\text{among mosquitoes in a collection}}$ = **0.31046**

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**Table 3.7 Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all four environmental terrain categories**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 4 environments | 5 | 3.563 | 0.01492 | 3.68 |
| Among collections in each environment | 36 | 16.869 | 0.15066 | **37.15***** |
| Among mosquitoes in a collection | 25 | 6.000 | 0.24000 | **59.17***** |
| Total | 66 | 26.433 | 0.40558 | |

Fixation Indices: $F_{\text{among land covers}}$ = 0.03679
***$F_{\text{among collections in each land cover}}$ = **0.38566**
***$F_{\text{among mosquitoes in a collection}}$ = **0.40826**

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**Table 3.8: Analysis of molecular variance of groups of S sequences isolated from mosquitoes collected from all three years**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 3 years | 2 | 8.709 | 0.16669 | **36.75*** ** |
| Among collections in each year | 53 | 18.008 | 0.12017 | **26.50*** ** |
| Among mosquitoes in a collection | 27 | 4.500 | 0.16667 | **36.75*** ** |
| Total | 82 | 31.217 | 0.45353 | |

Fixation Indices: ***$F_{\text{among years}}$ = **0.36754**
***$F_{\text{among collections in each year}}$ = **0.41895**
***$F_{\text{among mosquitoes in a collection}}$ = **0.63251**

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

**Table 3.9: Analysis of molecular variance of groups of M sequences isolated from mosquitoes collected from all three years**

| Source of variation | Degrees of freedom | Sum of squares | Variance components | % variation |
|---|---|---|---|---|
| Between collections in the 3 years | 2 | 5.246 | 0.10983 | **24.44*** ** |
| Among collections in each year | 47 | 18.770 | 0.19744 | **43.93*** ** |
| Among mosquitoes in a collection | 17 | 2.417 | 0.14216 | **31.63*** ** |
| Total | 66 | 26.433 | 0.44942 | |

Fixation Indices: ***$F_{\text{among counties}}$ = **0.24438**
***$F_{\text{among collections in each county}}$ = **0.58139**
***$F_{\text{among mosquitoes in a collection}}$ = **0.68369**

*p < 0.05, ** p < 0.01, *** p <0.001 are indicative of significant sequence variation

24.44% of the variation observed among the years the samples were collected ($p < 0.001$) (Table 3.9). These results indicate that the collection year may be a determinant of the M segment haplotype.

**7. Chi-square test for independence of virus haplotype and collection year**

A chi-square test for independence was performed to further establish the relationship between the year the mosquito was collected and the genetic sequence. The null hypothesis was that there is no significant relationship between the collection year and the genetic sequence. The chi-square values for the S segment ($\chi^2 = 77.32$, $p \leq 0.0001$) and the M segment ($\chi^2 = 71.42$, $p \leq 0.0001$) were extremely significant, suggesting a strong association between the genetic sequence and the collection year. The phylogenetic tree clearly demonstrates this distribution (Figure 3.3 and 3.4) and a graphical representation of these data can be found in Figure 3.5.

**8. Detection of quasispecies in LACV-infected mosquitoes**

There was also evidence of the existence of quasispecies in some of the mosquitoes analyzed. In these cases, alignment of sequences of 3 to 5 clones amplified from a single mosquito showed 2 or 3 different viral haplotypes for each segment (Figure 3.6). This occurred in 28/81 of the samples (34.6%) in the S segment and 22/69 of the samples (31.9%) in the M segment (Table 3.10). In the remaining samples, there was a strong consensus sequence for a given segment with no variable sites.

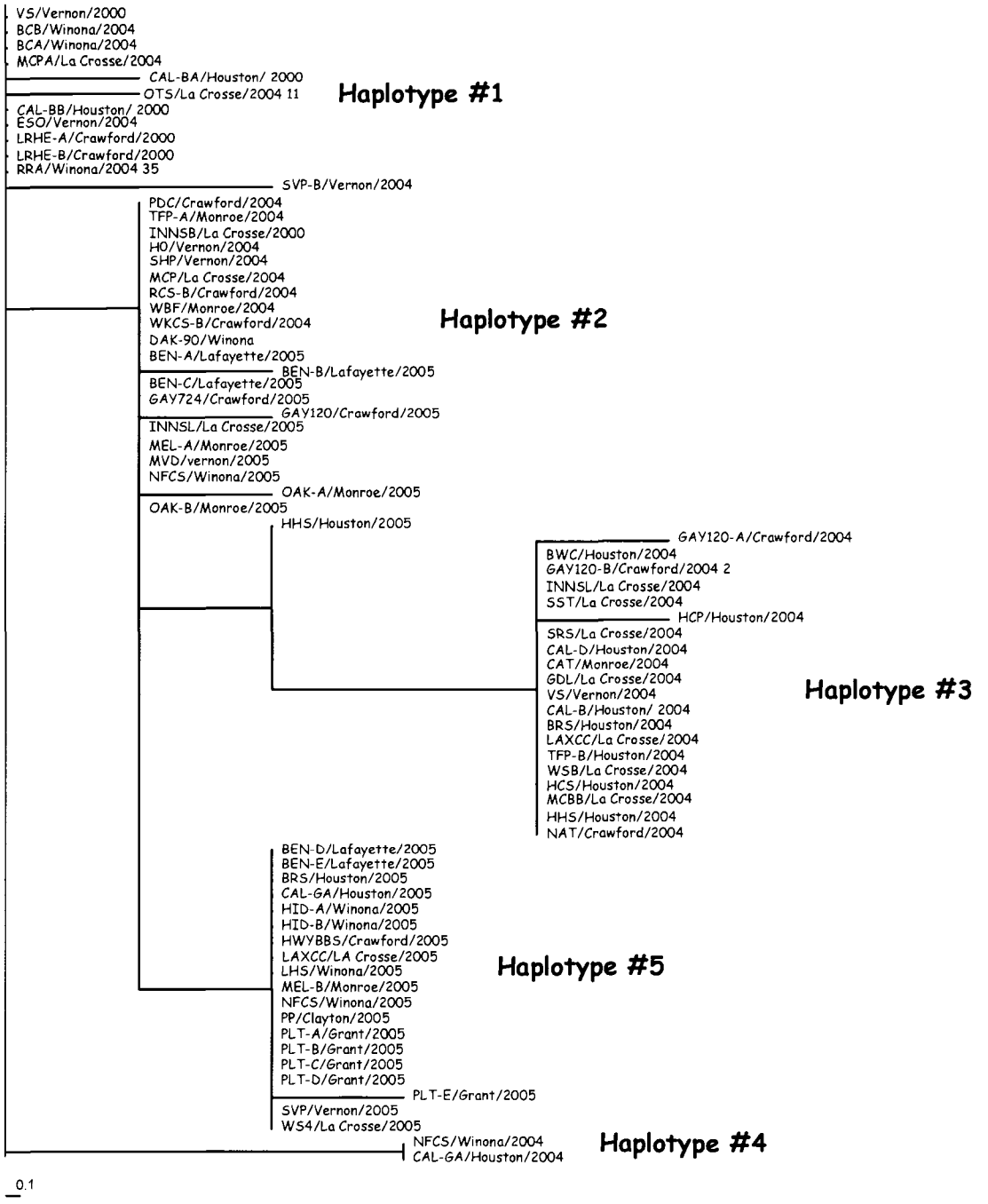



**Figure 3.3: S segment maximum parsimony tree reveals haplotypes segregating into clades by collection year.**

The three different collection years are distinguished by colors (blue is 2000, red is 2004, and green is 2005). Most of the LACV RNA amplified from mosquitoes collected in 2000 clustered in haplotype #1, LACV RNA amplified from mosquitoes collected in 2004 clustered in haplotypes 2-4, and many LACV RNA sequences amplified from mosquitoes collected from 2005 clustered into haplotype #5.

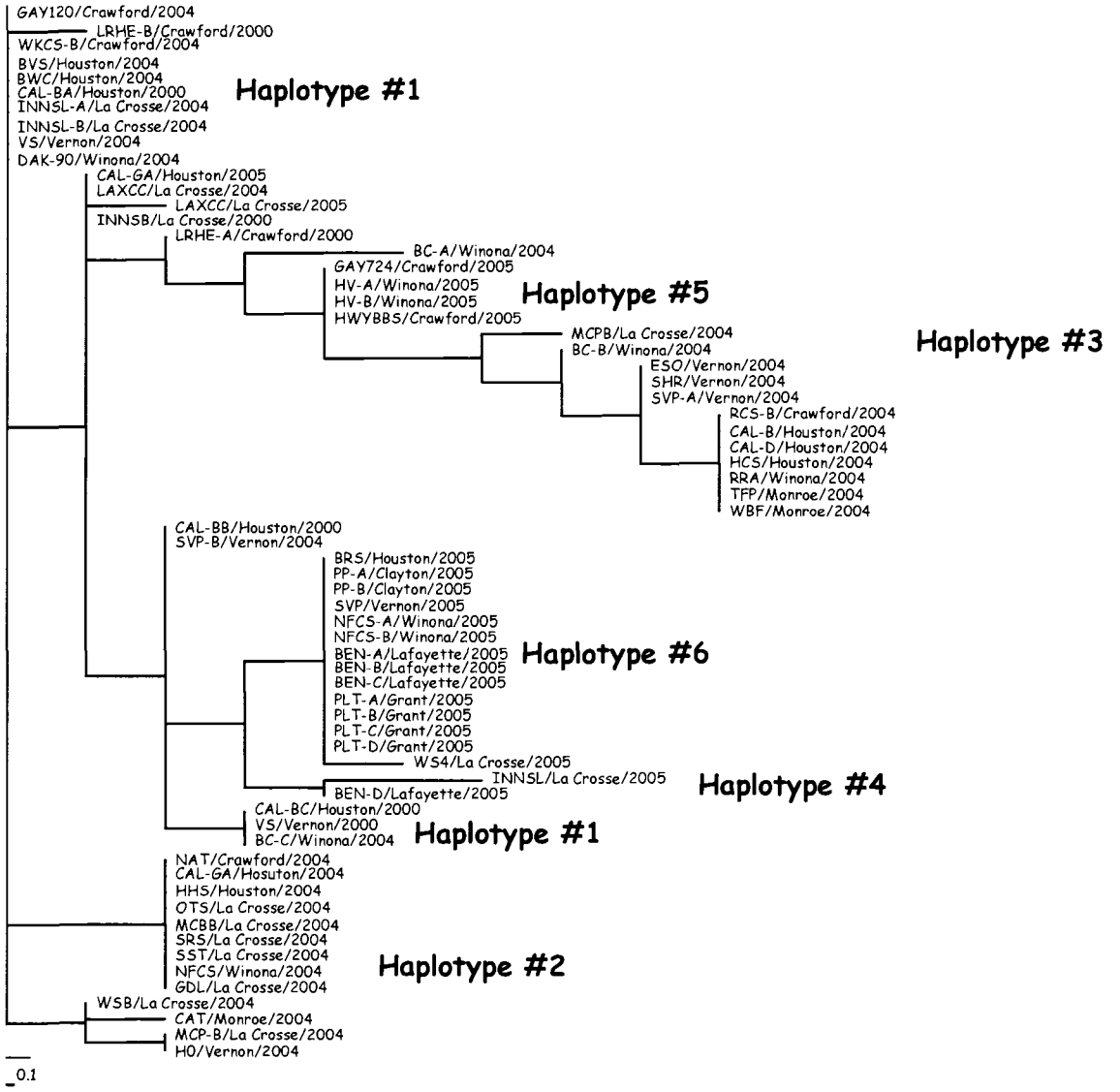



**Figure 3.4: M segment maximum parsimony tree revealed haplotypes cluster into clades by collection year.**

The three different collection years are distinguished by colors (blue is 2000, red is 2004, and green is 2005). All of the LACV RNA amplified from mosquitoes collected in 2000 clustered in haplotype #1, LACV RNA amplified from mosquitoes collected in 2004 clustered in haplotypes 2 and 3, and all LACV RNA amplified from mosquitoes collected in 2005 clustered into haplotypes 4-6.

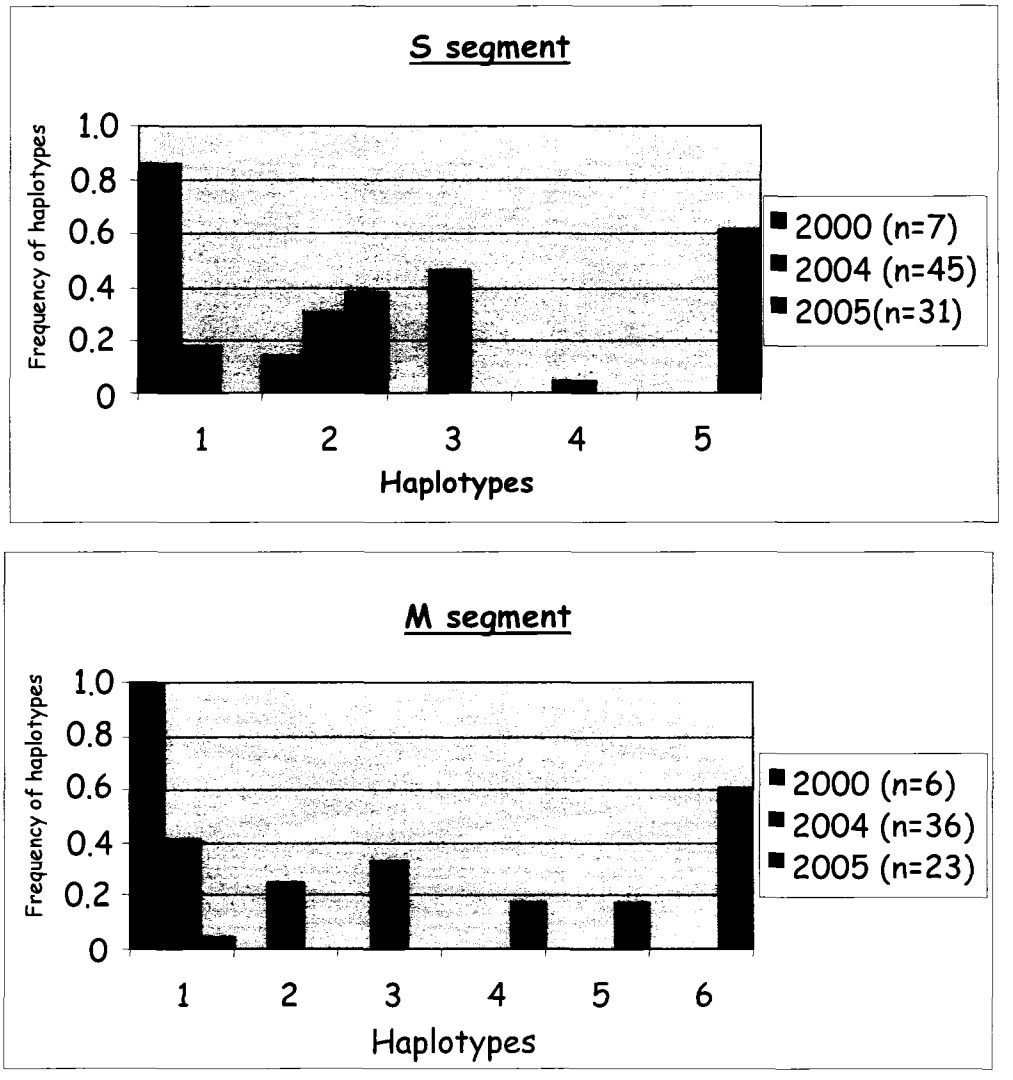



**Figure 3.5: The frequency of haplotypes is significantly associated with the collection year.**

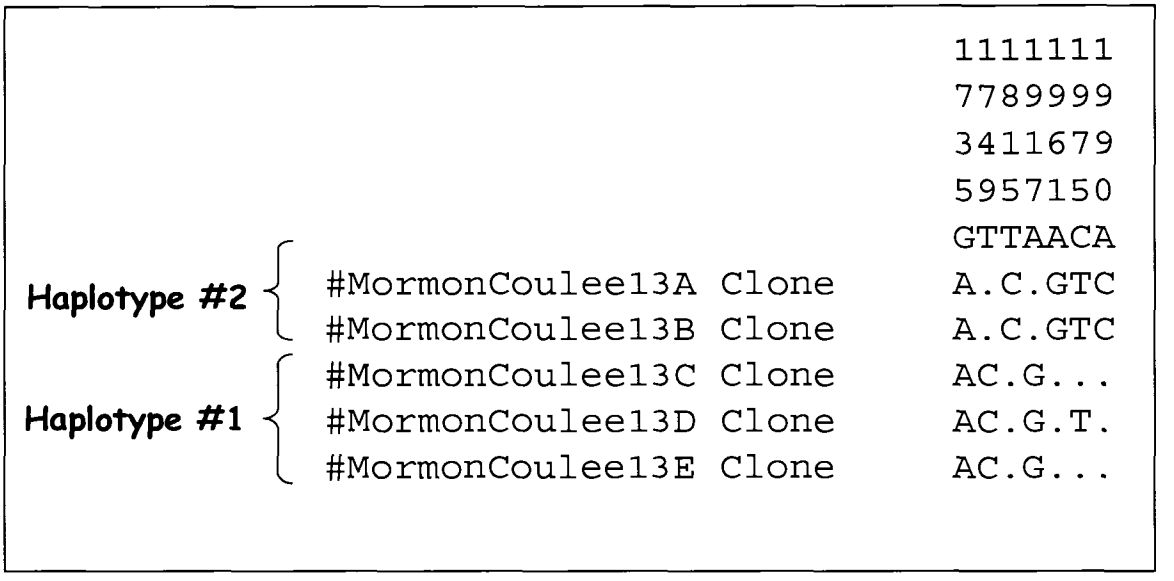



**Figure 3.6: Evidence for multiple virus haplotypes in a single mosquito**
The viral RNA sequence amplified from the mosquito collected from Mormon Coulee had viral sequence from both haplotypes 1 and 2 of the M segment when 5 clones were sequenced.

**Table 3.10: Evidence for quasispecies in LACV S and M segments in all 3 years LACV RNA was analyzed.**

| | | | |
|---|---|---|---|
| S | 2000 | 2 | 33.8 |
| S | 2004 | 14 | 31.8 |
| S | 2005 | 12 | 38.7 |
| M | 2000 | 0 | 0 |
| M | 2004 | 12 | 34.3 |
| M | 2005 | 10 | 34.5 |

## D. DISCUSSION

The studies suggest that there are no physical barriers to viral gene flow and no isolation by distance for LACV RNA sequences in the 15,360 $km^2$ study range. AMOVA tests investigating viral gene flow (grouped by quadrants, Table 3.4 and 3.5) and isolation by distance (grouped by counties, Table 3.2 and 3.3) demonstrated no significant geographical (spatial) determinant conditioning genetic variation of LACV. It should be noted that the AMOVA tests typically are performed on sample sizes of >25 and the sample size of RNA sequences analyzed was between 1 and 5 per collection site for this analysis, thereby inflating the results due to small sample size. Thus, only very significant AMOVA ($p \leq 0.001$) results were considered.

These results support previous data demonstrating significant genetic variation of LACV in the region (Klimas et al., 1981). Previous phylogeographic studies suggested that there is genetic variation of LACV in the U.S. with significant correlation between genotypes and geographic location. There are distinct lineages of LACV isolated in the upper Midwest, the southeast U.S. and the northeast U.S. (Armstrong & Andreadis, 2006, Klimas et al., 1981).

Since there is no isolation by distance and there are no barriers to viral gene flow that could prevent virus trafficking, all the S and M haplotypes can be expected to circulate throughout the 15,360 $km^2$ range of the study area. This could have epidemiological implications because there may be a narrow range of genotypes correlated with fatal outcomes (Huang et al., 1997). It is important to understand the potential movement of such genotypes. This indicates that more or less virulent LACV could traffic and be transmitted throughout the entire 15,360 $km^2$ range of the study.

An additional analysis was performed exploring the effect of the environmental and ecological factors on the haplotype. The LACV haplotypes could be determined by environmental factors including land cover, primary vertebrate hosts, and water sources. The different terrains examined were patchy landscape within 0.5 km distance of a small water source, mixed forest within a 1 km distance of the Mississippi River, residential/ urban area, and mixed forest without a water source (Table 3.1). The AMOVA analysis for the M segment revealed no significant variation in haplotypes between the different environmental terrains (Table 3.7). There was significant variation for the S segment observed in haplotypes between the environmental groups (Table 3.6) (11.28%, $p \leq 0.01$). Given that p was not $\leq 0.001$, this result was probably not significant and the results were inflated due to small sample sizes.

The association of environmental factors with the S segment haplotypes but not the M segment haplotypes is difficult to explain. There may be contributing factors that could produce genetic variability in the S segment as a result of interaction with specific land cover and ecological factors however the primary role of NSs is still being elucidated. The NSs protein has been identified as a virulence factor necessary to suppress production of interferon in infected cells and in Rift Valley fever virus infected animals (Weber & Elliot, 2002). NSs is thought to play a role in evading the innate immune response of the vertebrate hosts (Bridgen et al., 2001, Kohl et al., 2003, Soldan et al., 2005, Streitenfeld et al., 2003, Weber et al., 2002). The interaction between different vertebrate hosts and NSs could result in selection and therefore genetic variation. This is highly speculative, especially since a similar observation was not seen with the M segment, which conditions virus attachment to host cells and virulence.

Because the analysis revealed no significant viral RNA variation attributable to spatial or environmental factors, temporal correlates of viral genotypes were investigated. The AMOVA test revealed that both the S (36.75%, $p \leq 0.001$) and M (24.44%, $p \leq 0.001$) segment haplotypes clustered by collection year (Table 3.8 and 3.9). The data were then further analyzed with a Chi-square test for independence to determine the significance of this correlation. The S segment haplotypes ($\chi 2 = 77.32$, $p \leq 0.001$) and M segment haplotypes ($\chi 2 = 71.41$, $p \leq 0.0001$) were significantly correlated with the collection year (Figure 3.5).

The observed correlation between collection year and LACV haplotype could be due to using different collection sites each year. There were only ten collection sites of 48 total sites with mosquitoes analyzed in multiple years. However, the LACV haplotypes amplified from the mosquitoes collected from the ten collection sites used in multiple years strictly cluster with other haplotypes amplified from mosquitoes collected in the same year. For example, in the M segment phylogenetic analysis, BvilleRamsey2004 clusters into haplotype #1 with other 2004 samples, whereas BvilleRamsey2005 clusters into haplotype #6 with other 2005 samples. Mosquitoes collected from different collection sites were analyzed each year to increase the geographic distribution.

This correlation was surprising for a variety of reasons. Previous studies in the laboratory have suggested that LACV is genetically stable over time. RNA oligonucleotide fingerprinting analyses revealed that LACV genome remains stable during transovarial transmission (TOT) in the insect host and during transfer between the insect and vertebrate host (Baldridge et al., 1989). Furthermore, there were no nucleotide changes in the viral genome occurring during two generations of TOT in the mosquito

host. There was also no variation observed upon transmission to and replication in a vertebrate host by bite of a second generation transovarially infected mosquito. These results suggest that the virus does not appear to undergo rapid evolution during TOT in the mosquito host or during horizontal transmission from mosquito host to vertebrate host in laboratory conditions.

The genome sequence of low passage LACV isolates from 1960 (LACV/human/1960) and 1978 (LACV/mosquito/1978, LACV/human/1978) were analyzed for genetic variation over time (Bennett et al., 2007). The results indicated that circulating LACV are genetically stable over time and geographic distance with 97.8% to 100% amino acid identity for the N, NSs, M and L proteins. A phylogenetic analysis of a 1.6 kb portion of the M segment of LACV isolates from 1960-2005 revealed that the virus isolates from the upper Midwest of the U.S. between 1960 and 1981 were in the same phylogenetic clade (Armstrong & Andreadis, 2006). The isolates from Georgia, Missouri, New York, North Carolina, and Tennessee between 1974-1997 fell into a different clade, and the isolate from 2005 recovered in Connecticut fell into a third clade. This evidence clearly supports the hypothesis that LACV RNA is relatively stable over time and local geographic regions (Armstrong & Andreadis, 2006).

The results obtained in the current study do not support previous laboratory studies suggesting stability of LACV over time (Baldridge et al., 1989). One possible explanation for the genomic stability seen in laboratory LACV isolates could be multiple passages in mammalian cell cultures, mouse brains, or C636 cell cultures (*Aedes albopictus* cells). Passages of LACV *in vitro* could result in selection for more fit viruses in laboratory systems. This eliminates the selection forces imposed on viruses in the field

where the viruses must interact with various tissues in both vertebrate and invertebrate hosts and their respective immune responses. The LACV RNA analyzed in this study was viral RNA directly extracted from field-infected mosquitoes. The high amount of variability observed in this RNA may better reflect LACV evolutionary potential in nature.

The temporal variation observed within the M segment in the G1 glycoprotein gene suggests that diversifying selection due to immune factors, specifically neutralizing antibodies from the vertebrate host, possibly selects for new haplotypes each year (Figure 3.4 and 3.5). Previous evidence has demonstrated that LACV genome is responsive to selective pressure (Gonzalez-Scarano et al., 1987, Sundin et al., 1987). Monoclonal antibody-selected variants exhibited decreased cell fusion ability, loss of neuroinvasiveness in mice, and decreased infectivity in mosquitoes. These phenotypes then reverted to wildtype following passage in an infected mosquito and serial passage in BHK-21 cells.

The variation observed within the S segment by collection year (Figure 3.3 and 3.5) is more difficult to explain, due to minimal exposure to selective immunological forces. One possible explanation involves linkage between gene products. In segmented viruses all segments contain sequences important for packaging in both the non-coding region and the open reading frame (Fujii et al., 2005, Fujii et al., 2003, Muramoto et al., 2006, Watanabe et al., 2003). These sequences might serve as linking sites enhancing and facilitating the interaction with other vRNA segments. One segment can affect the incorporation of other segments, with the packaging efficiency being dependent on all segments present (Muramoto et al., 2006). The three LACV RNA segments complex

with the nucleoprotein and RNA polymerase and then these RNPs are packaged into progeny particles via an interaction between the nucleoprotein and the glycoproteins (Overby et al., 2007). The G1 and G2 glycoproteins line up on the membrane and N has to be recognized and packaged. The temporal determinants observed in the S segment sequences may be explained by linkage equilibrium due to the interaction between the nucleoprotein and the glycoproteins during packaging.

There was evidence for the presence of quasispecies in LACV RNA isolated from mosquitoes (Figure 3.6). Quasispecies yield multiple haplotypes that are available for packaging which may have developed due to adaptive pressure in the mosquitoes. This allows LACV to adapt to changes in the host and the vector. Some haplotypes may have a greater fitness in one environment compared to a different environment and this could allow LACV to have a successful transmission cycle between vector and vertebrate host. Although each mosquito may contain two or three haplotypes of the LACV RNA, only certain haplotypes may result in infectious particles. The haplotype packaged depends on the environment and the replicative fitness. The presence of quasispecies may contribute to the variability observed in this analysis. In order to determine if infectious virus has genome segments demonstrating quasispecies-like sequence variability, viruses will need to be isolated from mosquitoes and plaque-purified clones will need to be sequenced. The viral sequence will confirm if the temporal association with genetic variation is indeed occurring in the field. However, passage *in vitro* may also reduce the observed variability.

This study revealed a significant temporal association for the genetic sequence of the LACV S and M RNA segments (for the years 2000, 2004, and 2005). The variation

observed within the M segment could be associated with diversifying selection imposed by the vertebrate host. The variation seen in the S segment could be linked to selective pressures of the M segment. In this scenario, the S segment must change to retain the ability of the encoded nucleocapsid protein to interact with the G1 and G2 glycoproteins during packaging. Further studies are needed to determine if the variation observed within the LACV-infected mosquitoes reflects the epidemic potential of the virus.

# IV. IDENTIFICATION AND CHARACTERIZATION OF LA CROSSE VIRUS ISOLATED FROM SUPER-INFECTED *AEDES TRISERATUS* MOSQUITOES FROM THE FIELD

## A. INTRODUCTION

La Crosse virus (LACV) is the leading cause of arboviral pediatric encephalitis in the United States (Rust et al., 1999). LACV is primarily distributed east of the Mississippi River with endemic areas in the upper Midwest and the southeastern United States. The distribution of LACV reflects the distribution of its primary vector, *Aedes triseriatus*, commonly known as the Eastern treehole mosquito. LACV is maintained in nature between *Ae. triseriatus* mosquitoes and chipmunks and tree squirrels. The transmission cycle of LACV is unique because the virus can be transovarially transmitted to the progeny (Watts et al., 1973). This is thought to be the overwintering mechanism of the virus since *Ae. triseriatus* mosquitoes overwinter as diapausing eggs (Pantuwatana et al., 1974, Watts et al., 1974).

In nature, the transovarial transmission (TOT) and filial infection rates are much lower than those observed under laboratory conditions. The TOT rate (percentage of infected females that transmit virus to at least one of their progeny) observed in the laboratory can reach as high as 98% and filial infection rate (percentage of infected progeny from a single LACV-infected female) can be as high as 70% (Miller et al., 1977). These rates suggest that LACV can persist four years or longer in the absence of horizontal transmission in vertebrate hosts (Miller et al., 1977). However, in one field study, LACV was found in only 10 of 1,698 larvae processed (field infection rate = 0.06%) (Beaty & Thompson, 1975). In this study, larvae were collected from mesh covered tree holes, which ensured that all collected mosquitoes originated from overwintering eggs and only 0.0%-16.7% of the larvae were infected (Beaty & Thompson, 1975). In another study, the prevalence of LACV in *Ae. triseriatus* larvae

(calculated as minimum field infection rate) from overwintered eggs in southwestern Wisconsin ranged from 0.003 to 0.006 (Lisitza et al., 1977). Clearly, there is a significant difference between the field (0.0%-16.0%) and laboratory LACV infection rates (70.0%).

In this regard, there is evidence that TO infection of *Ae. triseriatus* mosquitoes has negative fitness effects on overwintering eggs. Embryos from TO-LACV-infected and uninfected *Ae. triseriatus* mosquitoes were induced into diapause and shipped to Wisconsin in 1993 to overwinter in natural conditions (McGaw et al., 1998). Samples were returned to AIDL monthly and assayed for mortality, diapause, and infection. A comparison of mortality rates, diapause status, and infection rates revealed that LACV infection does affect survival; a greater proportion of uninfected eggs (93.7%) successfully overwintered compared to LACV-infected eggs (83.3%). The mortality seen in LACV-infected eggs occurred after embryos broke diapause in the spring (McGaw et al., 1998). There was no difference between infected and uninfected mosquito survivability when embryos were still in diapause. This suggests that diapause attenuates deleterious virus effects on embryos and that diapause intensity and duration could condition the efficiency of both vector and virus overwintering. Although there was an increase in mortality of LACV-infected eggs, most embryos (83.3%) successfully overwintered, contributing to the population the following year (McGaw et al., 1998).

Mathematical models have been used to investigate parameters that condition LACV persistence in nature. TOT is less than 100% efficient and horizontal transmission may be an inefficient mechanism to compensate for the erosion of LACV prevalence that occurs during vertical transmission in nature. Horizontally-infected females must produce enough infected eggs to compensate for the inefficiency of vertical transmission in order

to maintain a "stable virus" prevalence from year to year in the vector population (Defoliart, 1983). Four factors complicate horizontal amplification in nature (1) wasting of vector species bites on non-amplifier species of vertebrates, (2) wasting of otherwise infective vector bites on the immune vertebrates, (3) the low ratio of mosquitoes becoming orally infected per viremic vertebrate host and (4) the high mortality of orally-infected vectors that occurs between the infectious bloodmeal and the second subsequent gonotrophic cycle (Defoliart, 1983). Herd immunity is especially critical in this regard. LACV antibody prevalence rates in forested areas can exceed 90%; thus most mosquito bites on chipmunks and tree squirrels would be wasted on dead end hosts (Gauld et al., 1974).

Modeling of LACV transmission using infection rates determined in field studies suggested that horizontal amplification produces only 2% of the number of infected eggs needed to maintain a stable virus prevalence in overwintering mosquitoes. The model considered variables such as mean survival to life events, number of bites delivered, number of bites on amplifiers, number of viremias produced, number of orally-infected mosquitoes and number of infected eggs. Although the rates were slanted as heavily as possible in favor of virus survival, the model suggested that LACV would not persist in nature (Defoliart, 1983).

A model was also developed for a close relative of LACV, Keystone virus (KEY, Family: *Bunyaviridae*, Genus: *Orthobunyavirus*, Serogroup: California), (Fine & LeDuc, 1978) that included variables such as the vertical transmission rate, the effect of the virus upon vector fertility and survival, vector densities and distributions, the proportion of susceptible hosts in the vertebrate population, the attractiveness of different vertebrates to

the vector and vector survival rates. Through the incorporation of rates observed in nature, the Fine and LeDuc model suggests that the maternal vertical transmission rate must be at least 0.1 in order for KEY to be maintained by a combination of vertical transmission and vertebrate mediated amplification. Although there are differences between the life cycle of LACV and KEY virus, this model also suggests that LACV cannot persist in nature. The infection rates detected in field collected larvae are less than 0.1 (Beaty & Thompson, 1975, Lisitza et al., 1977, Watts et al., 1974), thus the Fine and LeDuc model confirmed the findings of the previous study (Defoliart, 1983). Both models used field relevant infection rates and both predicted that CAL viruses will not survive. It should be noted that modeling with filial infection rates obtained in laboratory studies also strongly suggests that arboviruses could not persist by TOT alone for more than a few generations (Fine, 1975). Obviously there must be some other mechanism(s) that conditions LACV persistence in its endemic foci.

A possible mechanism for LACV maintenance in nature could be a stabilized infection of *Ae. triseriatus*. This was first observed with Sigma virus and *Drosophila melanogaster* (Seecof, 1968). When female *D. melanogaster* fruit flies were infected with Sigma virus by inoculation, they developed a systemic infection, and a small proportion of the developing oocytes became infected. The progeny derived from these eggs were transovarially infected. When the virus was transmitted to a small percentage of the progeny by infecting the developing oocytes directly, the virus was in a "nonstabilized" state (Fleuriet, 1988). However, if the germarium of a female became infected, she transmitted the virus to nearly 100% of her progeny, which was designated as a "stabilized" state.

Once stabilized infection occurs in a percentage of females in a natural population, the prevalence of viral infection will be maintained indefinitely at that level as long as certain conditions are met: (1) virus infection does not affect fertility, fecundity, development or survival of the host either positively or negatively, (2) additional mosquitoes are not added to the infected pool by horizontal transmission and (3) any detrimental effects are balanced by an increase in the infection rate due to horizontal infection (Turell et al., 1982). As discussed, one of the difficulties of accepting TOT as the sole mechanism for arbovirus maintenance in nature has been the fact that infection rates in field collected larvae are quite low (Tesh, 1980). The concept of stabilized infections increases the chance of viral survival persistence. The virus could be maintained in nature at a fairly constant level by a relatively small number of females with stabilized infections, while the virus infection rate in the total arthropod population would remain low (Tesh & Shroyer, 1980).

Stabilized infection has been demonstrated with California encephalitis (CE) virus and *Aedes dorsalis* mosquitoes infected by intrathoracic inoculation (Turell et al., 1982). Three mosquito strains were selected for high rates of maternal vertical transmission. Females that were intrathoracically inoculated with CE virus transmitted the virus vertically to over 90% of their progeny after two generations of selection. Infected progeny in these subpopulations transmitted virus at similar rates through five generations. Thus, stabilized CE viral infections could occur in transovarially-infected progeny of a natural vector species. This supports the hypothesis that stabilized infections of *Ae. triseriatus* may play an important epidemiological role in the maintenance of LACV.

There has been very little work done with field-infected *Ae. triseriatus* mosquitoes to determine if stabilized LACV infections occur. In one previous study, 0.03% of *Ae. triseriatus* mosquitoes were found to be super-infected with LACV collected as eggs from the field (Beck, 2007). The super-infected mosquitoes had between $10^{3.62}$ and $10^{4.89}$ genome equivalents (GE) of LACV in their ovaries as compared to other LACV-infected mosquitoes, which had a mean ovarian GE of $10^{1.54}$. The GE in the ovaries of the super-infected mosquitoes was actually similar to the $10^{4.77}$ GE detected in the AIDL highly selected TOT-permissive laboratory strains. It is possible that these field collected mosquitoes would have filial infection rates of 50-70%, similar to the TOT-permissive laboratory strain. The super-infected mosquitoes may play an important role in persistence of LACV in nature through a stabilized infection mechanism.

The goal of this study was to investigate the potential of LACV maintenance in nature through stabilized infection of *Ae. triseriatus* mosquitoes and to analyze the subsequent LACV isolates from super-infected mosquitoes. Mosquito eggs were collected in the field, returned to AIDL, hatched, and progeny were screened by an immunofluorescence assay (IFA) for LACV infection. Mosquitoes with large amount of LACV antigen were identified as super-infected. LACV was isolated from the super-infected mosquitoes, the RNA was sequenced and characterized. A portion of each of the three RNA segments was amplified from the super-infected mosquitoes and normally-infected mosquitoes and compared to determine if LACV obtained from super-infected mosquitoes was genetically different.

## B. METHODS AND MATERIALS

### 1. Egg collection

*Aedes triseriatus* eggs were collected from five oviposition traps in each of 119 sites in Minnesota (n = 30), Wisconsin (n = 83) and Iowa (n = 6) in areas where LACV encephalitis cases occurred or areas that contained clusters of people judged by the La Crosse County Public Health Department to be at risk for LACV infection (e.g. wooded areas adjacent to houses with children, schools, or playgrounds). Mosquito eggs were collected between mid-June and October of 2007 in Crawford (15 sites), La Crosse (37 sites), Monroe (12 sites), Vernon (12 sites), Lafayette (2 sites), Pierce (1 site) and Iowa (4 sites) counties in Wisconsin , Winona (13 sites) and Houston (17 sites) counties in Minnesota and Clayton county (6 sites) in Iowa (Figure 4.1). The eggs were transported to the insectaries at the Arthropod-borne and Infectious Diseases Laboratory (AIDL) at Colorado State University (CSU), Fort Collins, CO. Upon receipt, the eggs were immediately hatched and progeny were reared to adults.

### 2. Immunofluorescence assay

Mosquito legs were removed, squashed onto acid-washed microscope slides, and fixed in acetone. They were assayed for LACV antigen by direct immunofluorescence assay (IFA) using a LACV-specific polyclonal antibody (Beaty & Thompson, 1975). Super-infected *Ae. triseriatus* are quite evident by IFA when compared to uninfected and normally LACV-infected mosquitoes (Figure 4.2). The slides were scored on a scale of 0-5, with 0: uninfected mosquitoes, 1-4: normally-infected (NI) mosquitoes and 5: super-infected (SI) mosquitoes.

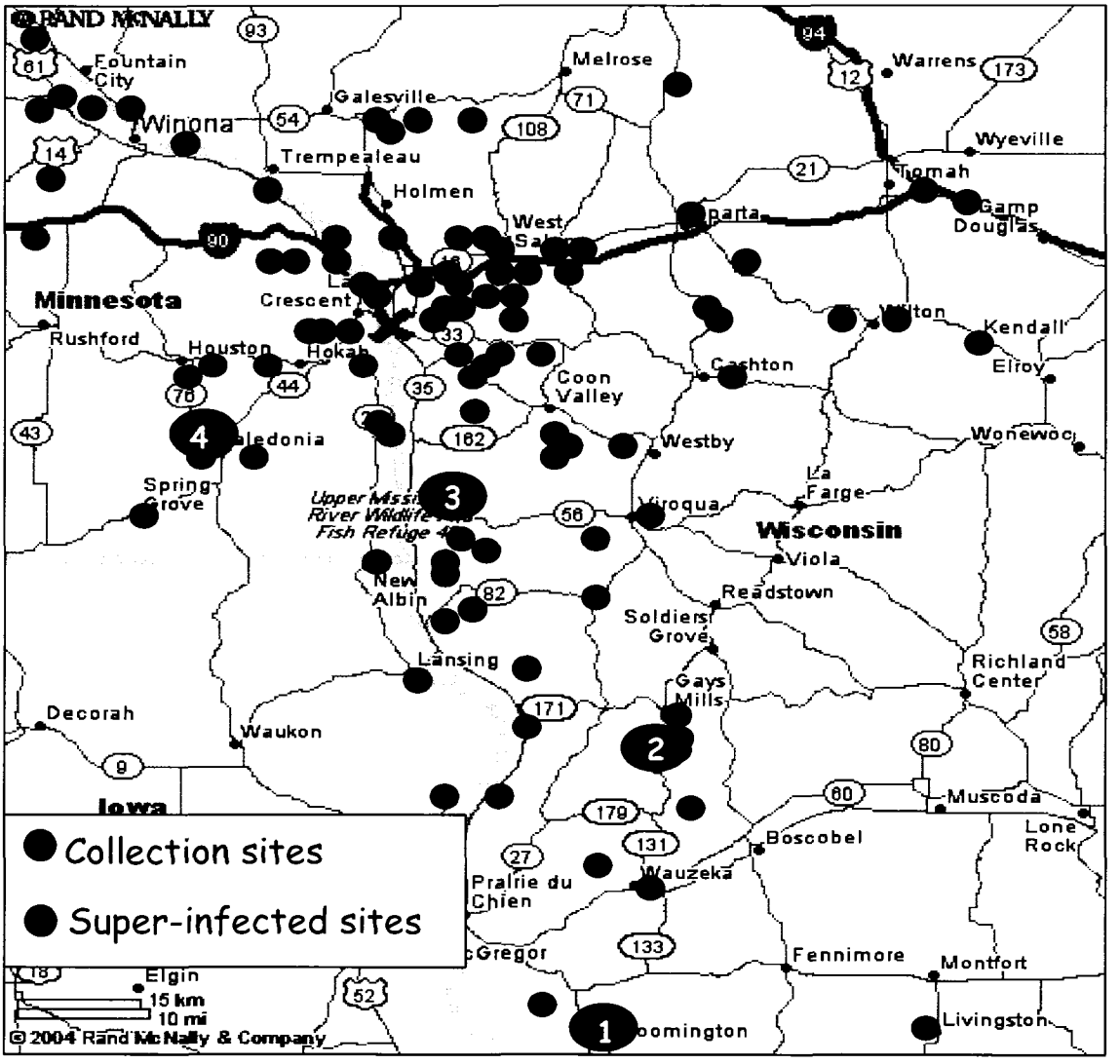



**Figure 4.1: Mosquito collection sites in Minnesota, Wisconsin and Iowa.**
Circles represent all collection sites. Red circles are the sites where LACV super-infected mosquitoes were collected in 2006 and 2007, Site 1: BEN2, Site 2: NAT, Site 3: SVP, and Site 4: CAL-GA.

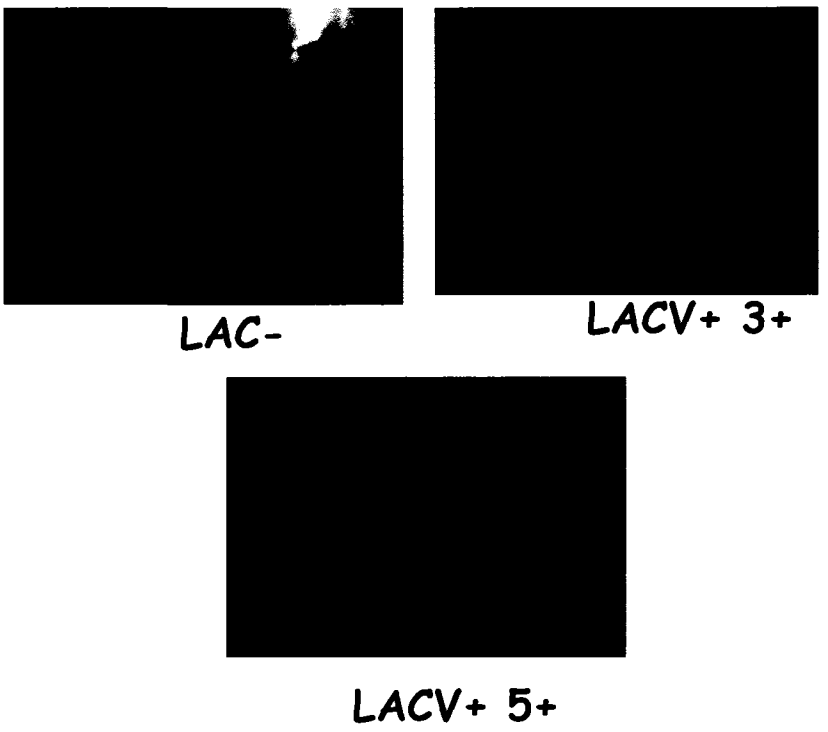



**Figure 4.2: IFA images from LACV-infected field mosquitoes**
The legs were removed from field-collected mosquitoes and tested for the presence of LACV antigen. The slides were scored on a scale of 0-5, 0, which is an uninfected mosquito, 3: a "normally infected mosquito" and 5: a "super-infected" mosquito.

**3. Virus isolation**

SI-mosquitoes, as identified by IFA, were triturated with a pellet pestle (Fisher Scientific) in a 1.5ml microcentrifuge tube containing 1ml of minimal essential medium (MEM) (Gibco), 2% fetal bovine serum, 200µg/ml penicillin/streptomycin, 200µg/ml fungicide, 7.1mM sodium bicarbonate, and 1x nonessential amino acids. The homogenate was centrifuged for 10 minutes at 500 x g to form a pellet.

Cell monolayers of Vero cells were grown in six-well plates at 37°C in an atmosphere of 5% $CO_2$. Supernatant from the centrifuged mosquito homogenate (0.2ml) was added to one well in a six-well plate, incubated at 37°C for one hour. Following the incubation, 5ml of medium (as described above) were added to each well. The presence of cytopathic effects (CPE) 5 days post-infection revealed the isolation of LACV (Gerhardt et al., 2001).

**4. $TCID_{50}$**

The $TCID_{50}$ of each SI-mosquito was determined by endpoint titration in Vero cells using microtiter plates (Scott et al., 1984). The mosquito homogenates were serially diluted $10^{-1}$ to $10^{-6}$ and 200µl of each dilution was added to one well of a 96 well plate. Five days post-infection, the endpoint was determined as the highest dilution with CPE. The virus titer was calculated by the Karber formula (Karber, 1931) and expressed as $\log_{10}$ $TCID_{50}$/ml.

**5. Plaque purification**

Monolayers of Vero cells in six-well plates were used for plaque purifying LACV isolates (Eckels et al., 1976). Virus isolates were serially diluted $10^{-1}$ to $10^{-6}$ and 200µl of each virus dilution was added to one well and incubated at 37°C for 1 hour. Following the

incubation, the virus inoculum was removed and 5ml of overlay (1g/100ml agar, 10x Medium 199, 10% fetal bovine serum, 7.1mM sodium bicarbonate, 0.2% diethylaminoethyl-dextran in Hank's BSS, 1x Eagle basal medium vitamins, 1x Eagle basal amino acids) were added to the well. The components of the overlay were combined at 44°C in the order listed to avoid solidification. After six days of incubation at 37°C in 5% $CO_2$, 200µl of the detection solution, methylthiazolyldiphenyl-tetrazolium bromide (MTT) (5mg/ml in PBS), was added to each well. The plates were incubated overnight and wells were examined for plaques. The plaques were individually picked and placed in 1ml of 0.2% MEM for 1hr at 37°C. The eluted virus was added to one well of a six-well plate. The cells were incubated for five to seven days and the presence of virus was confirmed by CPE.

**6. RNA purification from virus isolates and amplification by reverse transcription-PCR**

The supernatant and cells from the wells with plaque purified virus were removed and placed in a 15ml conical tube and centrifuged at 3000 rpm for 10 minutes. The supernatant was removed and the cell pellet was resuspended in 500µl of Trizol (Invitrogen). Total RNA was extracted according to manufacturer's instructions. The entire S, M, and L RNA segments were transcribed to cDNA using Superscript II reverse transcriptase (Invitrogen) and amplified by PCR using Ex Taq DNA polymerase (Takara) according to manufacturer's instructions. The entire S segment was amplified in two separate fragments, the M segment in three fragments, and the L in four fragments (Table 4.1). The reverse primers in Table 4.1 were used to synthesize the cDNA used for PCR. PCR was performed as follows: 94°C for 5 minutes, 37 cycles of [94°C for 1 minute,

57°C for 1 minute and 72°C for 2.5 minutes] followed by a final extension at 72°C for 8 minutes.

**Table 4.1: Primer sets for amplification of the three segments of LACV**

| | | | | |
|---|---|---|---|---|
| S-A | SF1 | AGTAGTGTACCCCACTTGAATAC | 1-23 | 525 |
| | SR | CTTAAGGCCTTCTTCAGGTATTGAG | 549-572 | |
| S-B | SF3 | CTTAAGGCCTTCTTCAGGTATTGAG | 453-476 | 486 |
| | SR1 | AGTAGTGTGCCCCACTGAATAC | 963-984 | |
| M-C | MF1 | AGTAGTGTACTACCAAGTATAGATGAACG | 1-27 | 2211 |
| | MR10 | GACTCCTTTCCTCTAGCAAGG | 2239-2258 | |
| M-D | MF9 | CAGACAACATGGAGAGTGTAC | 1798-1818 | 1786 |
| | MR5 | GTCAAATCTGGGAACTCCATTGCC | 3605-3628 | |
| M-E | MF15 | CAAGCTCATGGGGATGCGAAGAG | 3249-3271 | 1235 |
| | MR1 | AGTAGTGTGCTACCAAGTATA | 4507-4527 | |
| L-F | LF1 | AGTAGTGTACCCCTATCTACAAAAC | 1-25 | 2955 |
| | LR20 | GTTTTCCCTCTGTTCGCACTC | 2381-2401 | |
| L-G | LF8A | CAACTTGCCTACTATTCAAAC | 1899-1919 | 2164 |
| | LR7 | CCAATCCAACTGTACTAATCATTGAC | 4084-4108 | |
| L-H | LF8 | GCTACCAGGGCAGTCAAATGACCC | 3987-4011 | 1949 |
| | LR10 | CCTCTGCAACGTTAACTACACATACTG | 5961-5986 | |
| L-I | LF10 | CAGATATTGTCTGGTGGCCATAAAGCC | 5288-5314 | 1646 |
| | LR12 | AGTAGTGTGCCCCTATCTTC | 6961-6980 | |

**7. LACV RNA purification and amplification by reverse transcription-PCR from infected mosquitoes**

A portion of the mosquito homogenate (500µl) was used to amplify LACV RNA directly from an infected mosquito. Trizol (500 µl) was added to the homogenate, and RNA was extracted according to manufacturer's instructions. Portions of the S, M and L segments were amplified as described in II. B.4.

**8. Sequencing**

PCR products were separated by electrophoresis in 1% agarose gels with TAE buffer, stained with ethidium bromide, excised and extracted using a kit from Marligen Biosciences according to manufacturer's instructions. The PCR products were sequenced using the ABI PRISM dye terminator cycle sequencing kit (Applied Biosystems, CA) and the ABI 310 DNA automated sequencer at Macromolecular Resources, CSU.

**9. Maximum parsimony analysis**

Maximum parsimony phylogenetic analysis was performed using the Phylogenetic Analysis Using Parsimony (PAUP) 4.0b10 package (Swofford, 2003). The phylogenetic trees show the branches that appeared in the majority of the 1000 bootstrap pseudo replications and the frequency with which these appear among replications.

Two different maximum parsimony analyses were performed. One analysis consisted of comparison of LACV isolated from the SI-mosquitoes and previously published LACV sequences. LACV was isolated from five field-infected *Ae. triseriatus* mosquitoes and upon virus isolation RNA was extracted for RT-PCR to obtain sequence data. The entire S, M and L segments were sequenced and analyzed. For the remainder of

this chapter, these LACV sequences will be referred to as "virus isolates". The viruses used in this analysis can be found in Table 4.2.

A second maximum parsimony analysis consisted of comparison of LACV sequence data obtained from amplified LACV RNA from all the SI-mosquitoes collected in 2006 and 2007 and LACV RNA amplified from field-infected mosquitoes collected in 2000-2005 (Table 4.3). These sequences will be referred to as "LACV mosquito amplified sequences (MAS)" for the remainder of the chapter.

**10. Nucleotide sequence divergence between populations**

The virus isolates obtained from the SI-mosquitoes and the previously published sequences were divided into two different groups for the S, M and L segments; (1) virus isolates obtained from SI-mosquitoes and (2) previously published virus isolates. The nucleotide divergence was calculated between the populations as $d_{xy}$. The $d_{xy}$ is defined as the average number of differences between one sequence randomly chosen from population X and another sequence randomly chosen from population Y. This analysis can locate nucleotides that are specific in each population and possibly contributing to the super infection.

## C. RESULTS

**1. LACV-super-infected mosquito prevalence in the field**

The prevalence of SI-mosquitoes was determined in field collections from 2006 and 2007. In total, 19,471 mosquitoes collected in 2006 and 2007 were assayed by IFA. Of these, 740 (3.8%) were infected with LACV, but only 16 (0.02%) were SI. The LACV prevalence rate was 3.79% in 2006 compared to 3.81% in 2007 (Table 4.4). In 2006, 0.03% (2/6,149) were SI-mosquitoes compared to 0.01% (14/13,322) in 2007. In 2006,

**Table 4.2: Virus RNA sequences used in maximum parsimony analysis of virus isolates**

| | | | |
|---|---|---|---|
| Minnesota/Human/1960 | C6/36 2 | EF485030-EF485032 | (Bennett et al., 2007) |
| Alabama/Mosquito/1963 | Suckling mice 3 | DQ426682 | (Armstrong & Andreadis, 2006) |
| Ohio/Mosquito/1965 | Suckling mice 4, Vero cells 1 | DQ426683 | (Armstrong & Andreadis, 2006) |
| New York/Mosquito/1974 | Suckling mice 4, BHK 4 | D10370 | (Grady et al., 1987) |
| Wisconsin/Mosquito/1977 | unknown | DQ196118-DQ196120 | (Bennett et al., 2007) |
| Wisconsin/Human/1978-A | Mouse brain 1, BHK 2, Vero cells 1 | EF485033-EF485035 | (Bennett et al., 2007) |
| Wisconsin/Human/1978-B | Mouse brain 1, BHK2 | NC004108-NC004110[b] | |
| Rochester, MN/Mosquito/1978 | unknown | DQ426680 | (Armstrong & Andreadis, 2006) |
| DeSoto, WI/Human/1978 | Suckling mice 2, BHK 2 | U18980 | (Huang et al., 1995) |
| Richland County, WI/Mosquito/1978 | Suckling mice 2, BHK 1 | U70206 | (Huang et al., 1997) |
| North Carolina/Mosquito/1978-A | Mouse brain 1, Vero cells 3 | EF485036-EF485038 | (Armstrong & Andreadis, 2006) |
| North Carolina/Mosquito/1978-B | Suckling mice 1, Vero cells 2 | DQ426681 | (Armstrong & Andreadis, 2006) |
| Crawford County, WI/Mosquito/1979 | Suckling mice 2, BHK 1 | U70207 | (Huang et al., 1997) |
| Washington County, WI/Mosquito/1981 | Suckling mice 2, BHK 1 | U70208 | (Huang et al., 1997) |
| Georgia/Canine/1988 | Vero cells 1, Suckling mice 1 | DQ426684 | (Armstrong & Andreadis, 2006) |
| Missouri/Human/1993 | Vero cells 1 | U70205 | (Huang et al., 1997) |
| West Virginia/Mosquito/1995 | Vero cells 1 | DQ426685 | (Armstrong & Andreadis, 2006) |
| North Carolina/Mosquito/1997 | Vero cells 1 | DQ426686 | (Armstrong & Andreadis, 2006) |
| Tennessee/Mosquito/2000 | Vero cells 1 | DQ426687 | (Armstrong & Andreadis, 2006) |
| Connecticut/Mosquito/2005 | Vero cells 1 | DQ426688 | (Armstrong & Andreadis, 2006) |
| SVP/Vernon, WI/Mosquito/2006 | Vero cells 2 | | |
| NAT/Crawford, WI/Mosquito/2006 | Vero cells 2 | | |
| NAT/Crawford, WI/Mosquito/2007 | Vero cells 2 | | |
| BEN2/Lafayette, WI/Mosquito/2007 | Vero cells 2 | | |
| CAL-GA/Houston, MN/Mosquito/2007 | Vero cells 2 | | |

[a] All viruses isolated from mosquitoes were isolated from *Ae. triseriatus* except for AL/Mosquito/1963 (*Ps. howardii*) and TN/Mosquito/2000 (*Ae. albopictus*)
[b] Previous submission by Hughes *et al.* 2002.

**Table 4.3: MAS used in maximum parsimony analysis***

| | | | | | |
|---|---|---|---|---|---|
| LRHE-A/Crawford/2000 | La Crosse, WI | 2000 | | | x |
| VS/Vernon/2000 | Vernon, WI | 2000 | x | x | x |
| BWC/Houston/2004 | Houston, MN | 2004 | x | | |
| CAL-GA/Houston/2004 | Houston, MN | 2004 | x | x | x |
| DAK-90/Winona/2004 | Winona, MN | 2004 | x | | |
| MCPA/La Crosse/2004 | La Crosse, WI | 2004 | | x | |
| MCPB/La Crosse/2004 | La Crosse, WI | 2004 | | x | |
| NAT/Crawford/2004 | Crawford, WI | 2004 | x | x | x |
| NFCS/Winona/2004 | Winona, MN | 2004 | x | | |
| SVP/Vernon/2004 | Vernon, WI | 2004 | x | x | x |
| BEN2/Lafayette/2005 | Lafayette, WI | 2005 | x | x | x |
| BEN1/Lafayette/2005 | Lafayette, WI | 2005 | | x | |
| HHS/Houston/2005 | Houston, MN | 2005 | x | | |
| CAL-GA/Houston/2005 | Houston, MN | 2005 | x | x | x |
| HV/Winona/2005 | Winona, MN | 2005 | | x | |
| MVR/Vernon/2005 | Vernon, WI | 2005 | | x | |
| **NAT/Crawford/2006** | **Crawford, WI** | **2006** | **x** | **x** | **x** |
| **SVP/Vernon/2006** | **Vernon, WI** | **2006** | **x** | **x** | **x** |
| **BEN2/Lafayette/2007-C** | **Lafayette, WI** | **2007** | **x** | **x** | **x** |
| **BEN2/Lafayette/2007-D** | **Lafayette, WI** | **2007** | | **x** | **x** |
| **BEN2/Lafayette/2007-E** | **Lafayette, WI** | **2007** | **x** | **x** | **x** |
| **BEN2/Lafayette/2007-F** | **Lafayette, WI** | **2007** | **x** | **x** | **x** |
| **CAL-GA/Houston/2007-G** | **Houston, MN** | **2007** | **x** | **x** | **x** |
| **CAL-GA/Houston/2007-H** | **Houston, MN** | **2007** | **x** | **x** | **x** |
| **CAL-GA/Houston/2007-I** | **Houston, MN** | **2007** | **x** | **x** | **x** |
| **CAL-GA/Houston/2007-J** | **Houston, MN** | **2007** | **x** | **x** | **x** |
| **CAL-GA/Houston/2007-K** | **Houston, MN** | **2007** | **x** | **x** | **x** |
| **NAT/Crawford/2007-A** | **Crawford, WI** | **2007** | **x** | **x** | **x** |
| **NAT/Crawford/2007-B** | **Crawford, WI** | **2007** | **x** | **x** | **x** |

*The MAS were amplified from mosquitoes collected between 2000 and 2005. The MAS that are bolded were obtained from the SI-mosquitoes collected in 2006 and 2007. The S, M and L columns indicate sequences used in the analysis.

SI-mosquitoes were collected from Vernon County (SVP/Vernon, WI/Mosquito/2006) and Crawford County (NAT/Crawford, WI/Mosquito/2006). In 2007, SI-mosquitoes were collected from Crawford County (NAT/Crawford, WI/Mosquito/2007), Lafayette County (BEN2/Lafayette, WI/Mosquito/2007) and Houston County (CAL-GA/Houston, MN/Mosquito/2007) (Table 4.4).

The amount of LACV antigen detected in leg tissue of mosquitoes collected as eggs, hatched and reared to adults varied significantly. SI *Ae. triseriatus* mosquitoes contained extraordinary amounts of LACV antigen as evidenced by IFA. Uninfected mosquitoes contained no viral antigen and NI-mosquitoes contained very little LACV antigen (Figure 4.2). The majority (97.2%) of the LACV-infected mosquitoes were NI and scored 1-4; 2.76% of the LACV-infected mosquitoes were SI (Figure 4.3).

To further investigate the prevalence of SI-mosquitoes, multiple egg liners were hatched from "hot spots" where >1 SI-mosquito had been found. This resulted in detection of multiple SI-mosquitoes from these sites. In the NAT/2007 site, 0.84% (4/475) of the mosquitoes were SI (Table 4.5). In the BEN2/2007 site, 1.82% (4/220) were SI, and in the CAL-GA/2007 site, 12.1% (7/58) were SI-mosquitoes. The LACV prevalence rates for each site were compared to the overall prevalence rate of 3.9% using a Fisher's exact test to determine if the prevalence rates were significantly different. Results can be found in Table 4.5. The prevalence of LACV-infected mosquitoes in NAT/2007 and CAL-GA/2007 was significantly different ($p \leq 0.001$) than the overall prevalence rate of LACV-infected mosquitoes. The prevalence rates observed in SVP/2006 and BEN2/2007 were also significantly different ($p \leq 0.05$) than the overall prevalence rate of 3.9%. The only site whose prevalence rate was not significantly

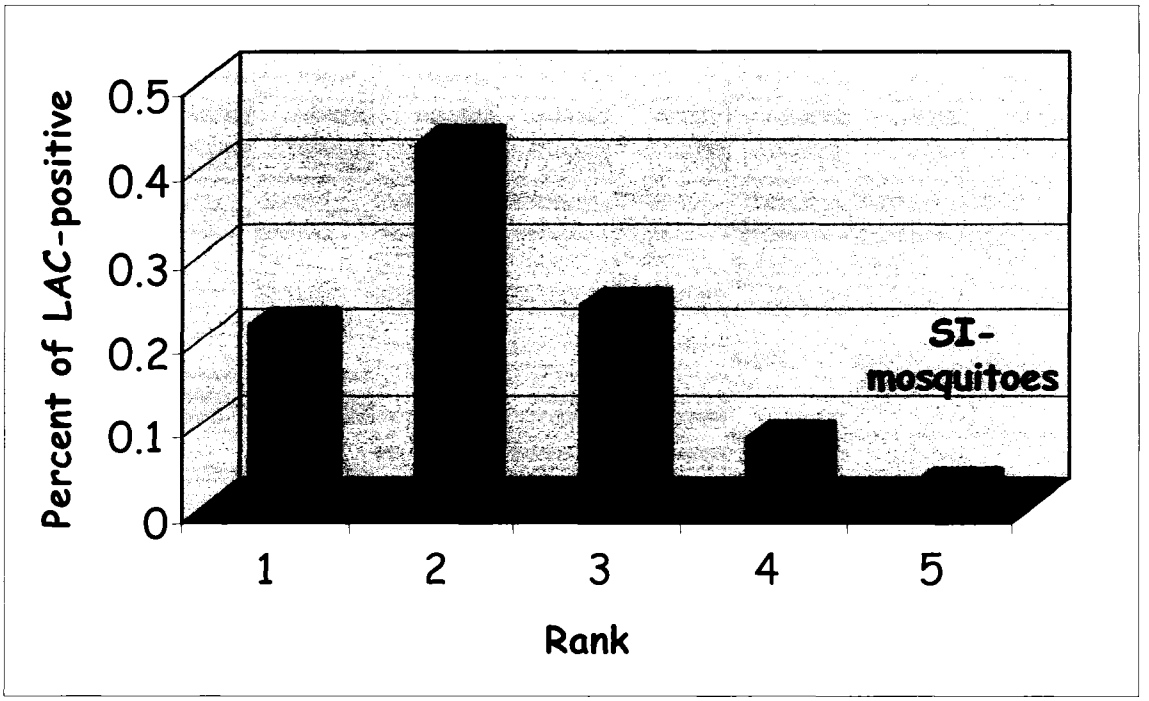



**Figure 4.3: The distribution of LACV-infected mosquitoes collected in 2007.** Mosquitoes were scored on a scale of 0-5 by the amount of detectable antigen by IFA. 1-4: NI-mosquitoes; 5: SI-mosquitoes.

different from the overall prevalence rate was NAT/2006, and this could be due to the small number of mosquitoes tested for LACV compared to the number of positive mosquitoes that were identified. Field collected mosquitoes were tested throughout the summer and interestingly, the SI-mosquitoes were only identified in each collection site once a year. The dates when the SI-mosquitoes were collected date is provided in Table 4.5.

**Table 4.4: Prevalence and distribution of LACV normally-infected and LACV super-infected mosquitoes in the 2006 and 2007 collections**

| | | | | |
|---|---|---|---|---|
| Clayton, IA | N/A | 1.50% | 0 | 0 |
| Crawford, WI | 2.16% | 4.40% | 0.180% | 0.025% |
| Iowa, WI | 2.78% | 2.12% | 0 | 0 |
| La Crosse, WI | 1.98% | 4.10% | 0 | 0 |
| Lafayette, WI | 3.88% | 6.90% | 0 | 1.100% |
| Monroe, WI | 0.55% | 4.20% | 0 | 0 |
| Vernon, WI | 3.30% | 2.54% | 0.089% | 0 |
| Houston, MN | 5.10% | 3.68% | 0 | 0.053% |
| Winona, MN | 2.80% | 3.81% | 0 | 0 |
| | | | | |
| Total mosquitoes tested | 6149 | 13322 | 6149 | 13322 |
| Total LACV+ or SI | 233 | 507 | 2 | 14 |
| Overall prevalence | 3.79% | 3.81% | 0.033% | 0.011% |

**Table 4.5: Prevalence of super-infected mosquitoes by collection site in 2006 and 2007**

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| SVP/2006 | Vernon, WI | 8/31/2006 | 84 | 7 | 8.30% | * | 1 | 1.20% |
| NAT/2006 | Crawford, WI | 2006 | 67 | 5 | 7.46% | | 1 | 1.50% |
| Nat/2007 | Crawford, WI | 7/17/2007 | 475 | 30 | 6.30% | *** | 4 | 0.84% |
| BEN2/2007 | Lafayette, WI | 9/10/2007 | 220 | 15 | 6.82% | * | 4 | 1.82% |
| CAL-GA/2007 | Houston, MN | 8/27/2007 | 58 | 7 | 12.10% | *** | 7 | 12.10% |

* $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$

**2. LACV titer of super-infected mosquitoes**

The LACV titer of the abdomen of the SI-mosquitoes ranged from 2.70 to 4.70 $\log_{10}TCID_{50}$/ml (Table 4.6). This is not significantly different from the average titer

observed in the TOT-permissive laboratory colonized mosquitoes (3.93 $\log_{10}TCID_{50}$/ml) (t-test: -1.96, $p > 0.05$).

**Table 4.6: LACV titer of super field-infected mosquitoes**

| | |
|---|---|
| NAT/Crawford/2006 | 3.03 |
| SVP/Vernon/2006 | 3.53 |
| NAT/Crawford/2007-A | 3.03 |
| NAT/Crawford/2007-B | 2.70 |
| BEN2/Lafayette/2007-C | 4.70 |
| BEN2/Lafayette/2007-D | 3.03 |
| BEN2/Lafayette/2007-E | 3.53 |
| BEN2/Lafayette/2007-F | 2.70 |
| CAL-GA/Houston/2007-G | 3.53 |
| CAL-GA/Houston/2007-J | 2.87 |
| CAL-GA/Houston/2007-K | 3.03 |
| TOT-permissive strain A | 3.93 |

**3. Phylogenetic analysis of virus isolates from SI-mosquitoes**

Viruses were isolated from each of the SI-mosquitoes. Following elution from purified plaques, RNA was extracted, sequenced and analyzed. The virus isolates were passaged two times in Vero cells. One virus isolate was analyzed from each collection site (SVP/2006, NAT/2006, NAT/2007, CAL-GA/2007, and BEN2/2007). Maximum parsimony trees were created for the sequences of the entire S, M and L segments (Figures 4.4-4.6).

The S segment maximum parsimony tree (Figure 4.4) produced four different clades; three clades containing virus isolates from Wisconsin and Minnesota and one clade containing the isolate from North Carolina and the isolates from NAT/2006 and NAT/2007. The other three LACV isolates from the SI-mosquitoes (CAL-GA/2007, BEN2/2007 and SVP/2006) are found in the clade with isolates from Wisconsin and

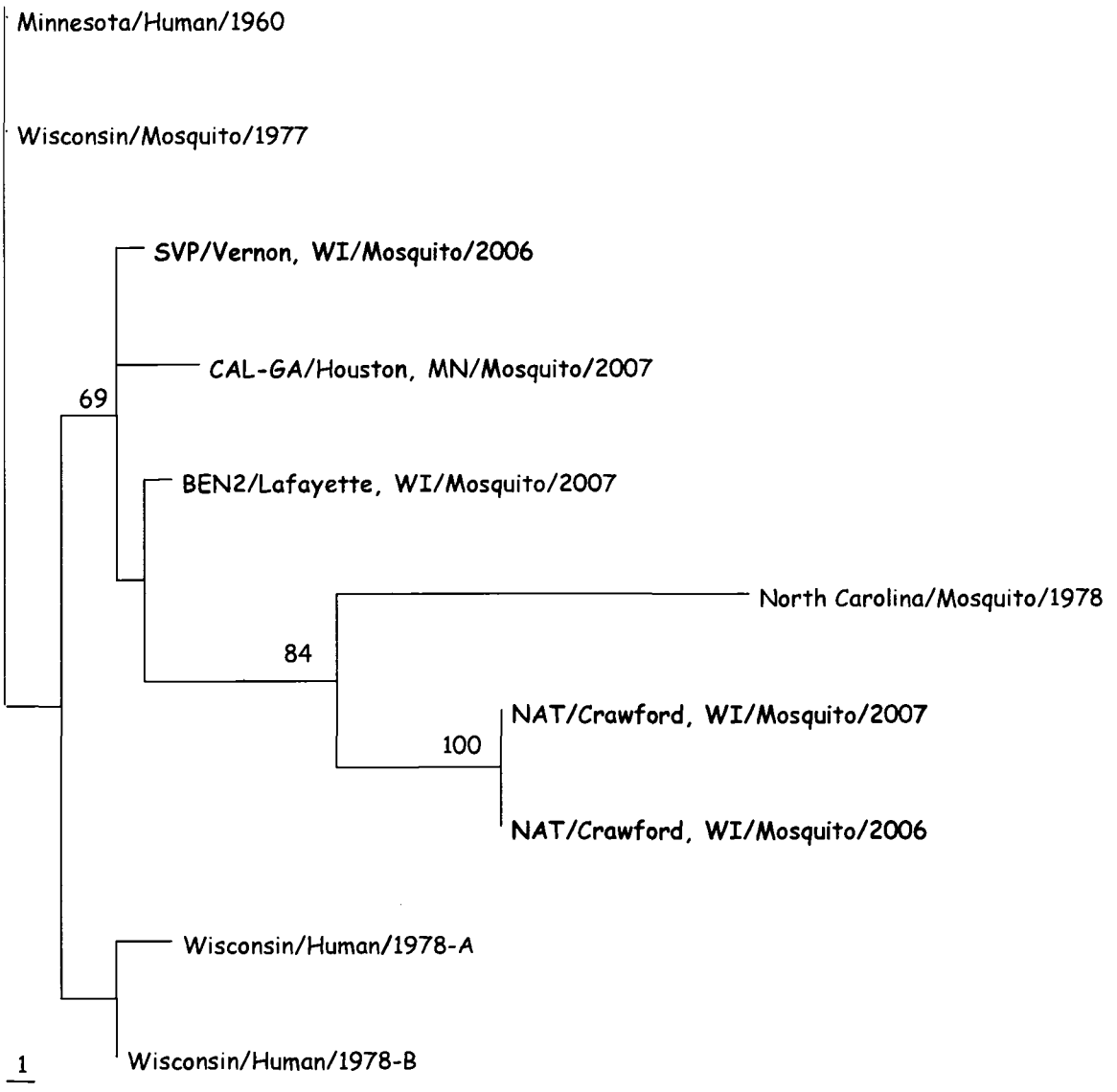



**Figure 4.4: Maximum parsimony phylogenetic analysis of the entire S segment of LACV isolates**

Phylogenetic analysis of the complete LACV S segment sequences from isolates obtained from SI-mosquitoes collected in 2006 and 2007 and previously published LACV S segment sequences. Bootstrap values were assigned for 1000 replicates and indicated by the numbers on the branches. Colors represent the collection sites of the SI-mosquitoes (Purple: SVP/2006, Red: BEN2/2007, Green: NAT/2006 and NAT/2007, and Blue: CAL-GA/2007).

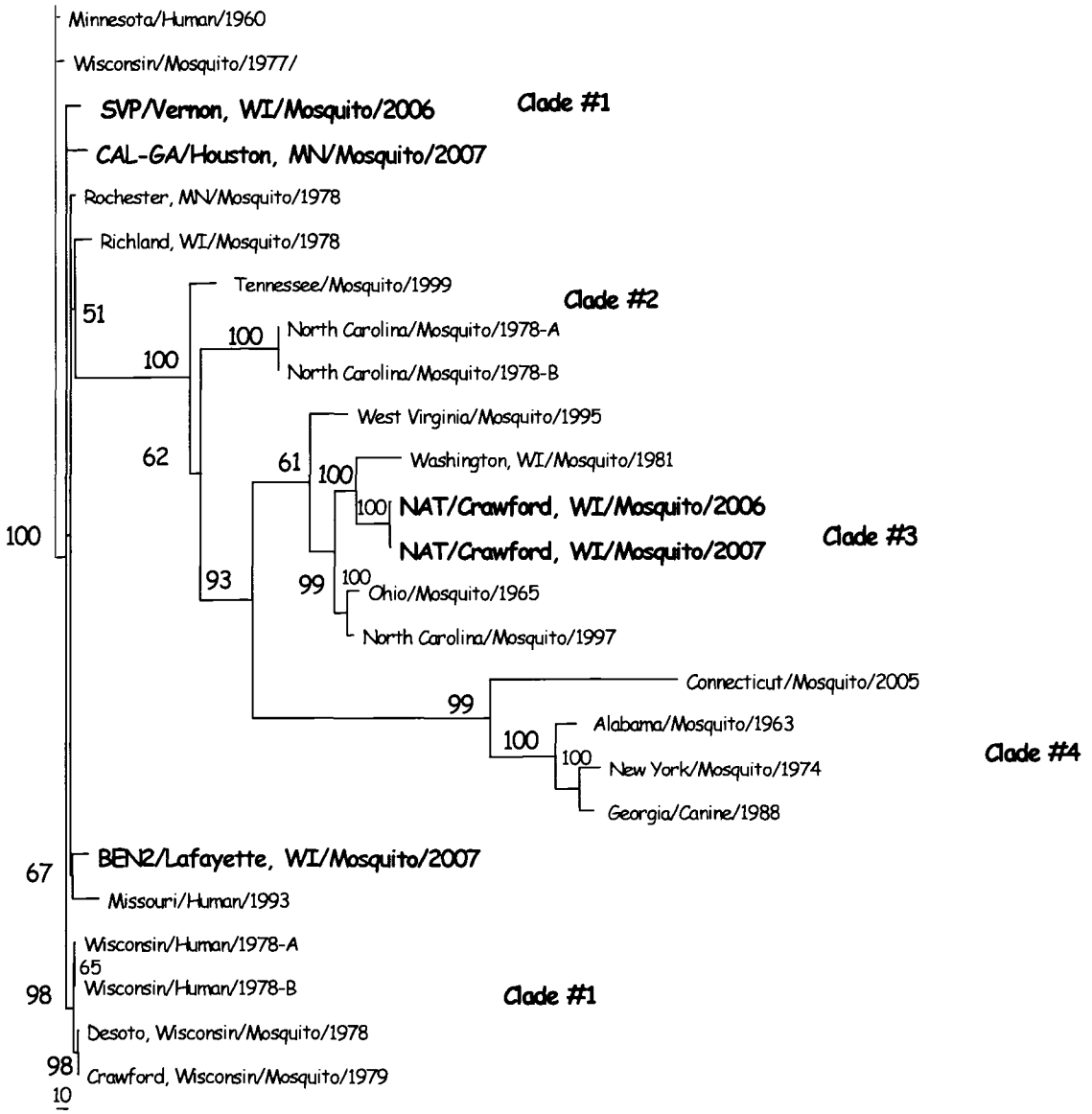



**Figure 4.5: Maximum parsimony phylogenetic analysis of the entire M segment of LACV isolates**

Phylogenetic analysis of the entire LACV M segment sequences from viruses isolated from SI-mosquitoes collected in 2006 and 2007, and previously published LACV M segment sequences. Bootstrap values were assigned for 1000 replicates and are indicated by the numbers on the branches. Colors represent the collection sites of the SI-mosquitoes (Purple: SVP/2006, Red: BEN2/2007, Green: NAT/2006 and NAT/2007, and Blue: CAL-GA/2007).

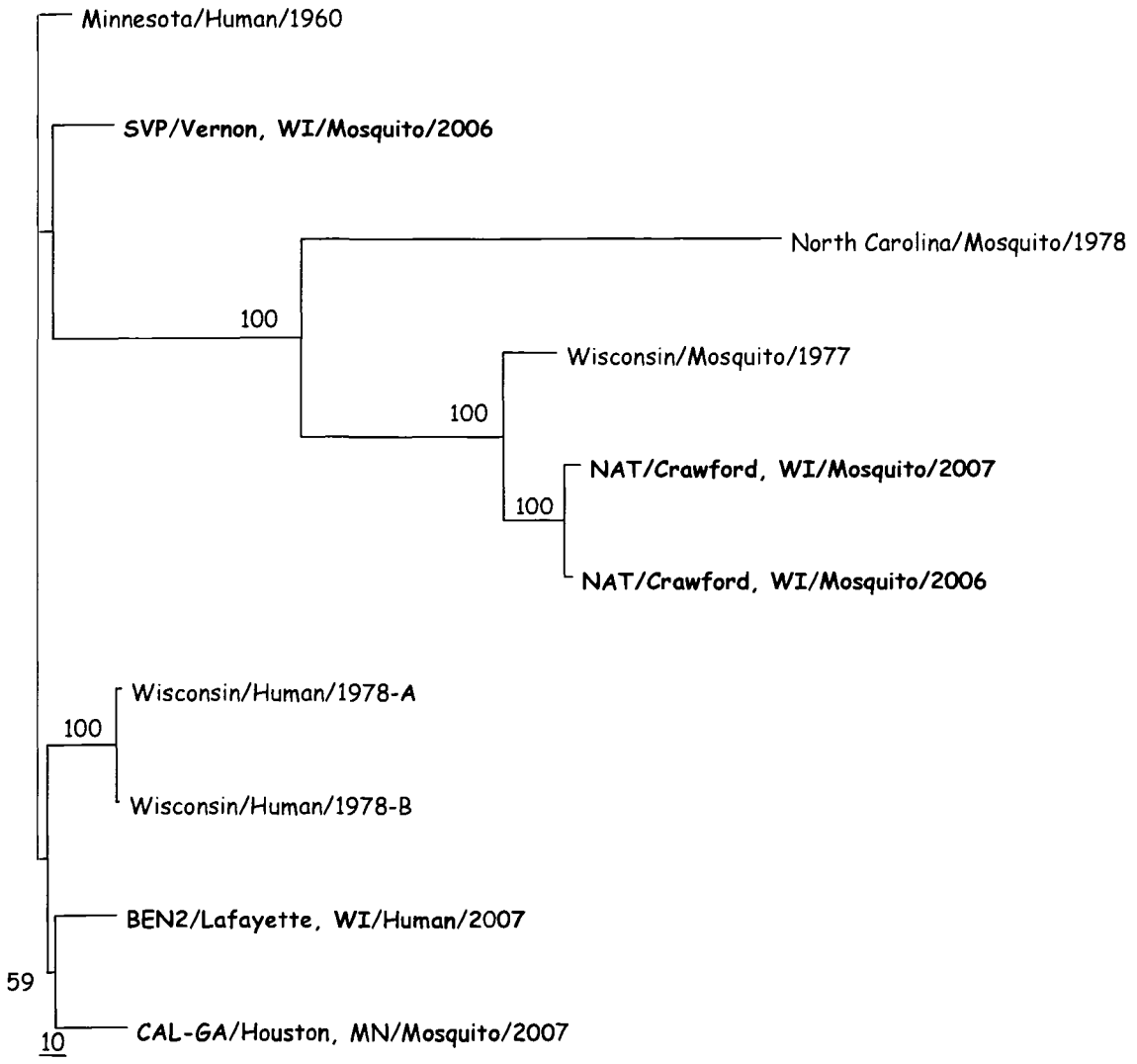



**Figure 4.6: Maximum parsimony phylogenetic analysis of the entire L segment of LACV isolates**

Phylogenetic analysis of the entire LACV L segment sequences from viruses isolated from SI-mosquitoes collected in 2006 and 2007, and previously published L segment sequences. Bootstrap values were assigned for 1000 replicates and are represented by the numbers on the branches. Colors represent the collection sites of the SI-mosquitoes (Purple: SVP/2006, Red: BEN2/2007, Green: NAT/2006 and NAT/2007, and Blue: CAL-GA/2007).

Minnesota. The phylogenetic analysis suggests that the M segment is more variable than the S segment (Figure 4.5). There are also more available published M segment sequences. The tree is distributed into four different clades associating somewhat with geographic distribution. Clade #2 contains virus isolates from Tennessee and North Carolina (1978), and clade #4 contains virus isolates from Connecticut, Alabama, New York, and Georgia. Clades #1 and #3 contain virus isolates recovered primarily from the Wisconsin and Minnesota region with the exception of West Virginia, Ohio, and North Carolina (clade #3) and Missouri (clade #1). As with the S segment, the virus isolates recovered from SVP/2006, BEN2/2007 and CAL-GA/2007 are found in the clade with most of the isolates recovered from Wisconsin and Minnesota. The isolates recovered from NAT/2006 and NAT/2007 are in the clade with the isolates recovered from North Carolina (1997), Ohio, and West Virginia and one isolate recovered from Washington County, WI.

The L segment maximum parsimony tree does not contain as many viruses because of the small number of available published sequences (Figure 4.6). The tree morphology is similar to the S segment. One very interesting observation is the location of the isolate Wisconsin/Mosquito/1977. In the S segment analysis, it is in the clade with all the LACV isolates from Wisconsin and Minnesota, whereas in the L segment analysis the isolate is in the clade with the isolates from NAT/2006 and NAT/2007. This is possible evidence for segment reassortment.

**4. Phylogenetic analysis of MAS from SI-mosquitoes and NI-mosquitoes**

Viral RNA was amplified directly from 11 SI-mosquitoes collected from SVP/2006, NAT/2006, NAT/2007, BEN2/2007, and CAL-GA/2007 for phylogenetic

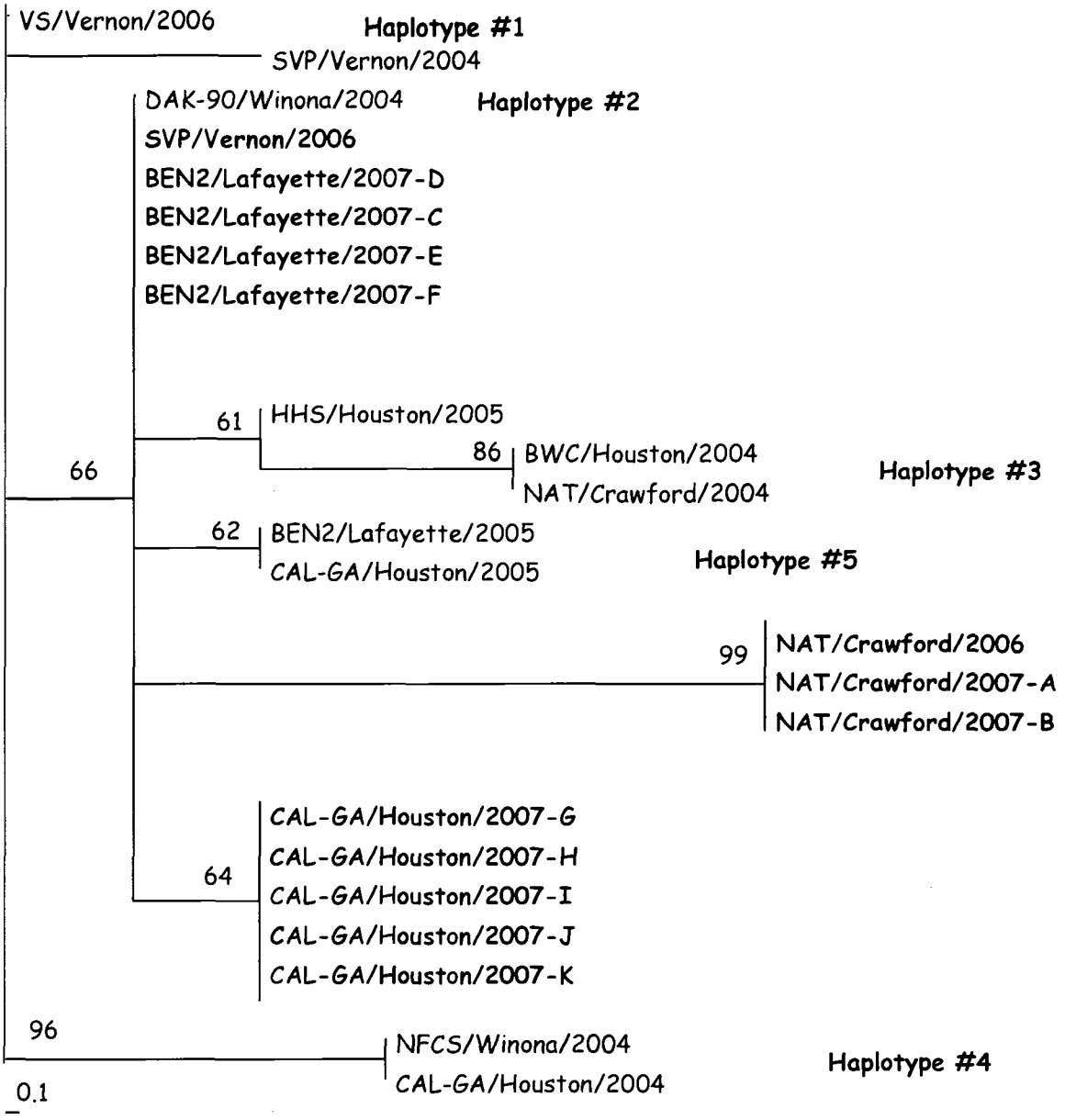



**Figure 4.7: Maximum parsimony phylogenetic analysis of a portion of the S segment RNA of MAS from normally-infected and super-infected field collected mosquitoes**
Phylogenetic analysis of MAS from field SI-mosquitoes collected in 2006 and 2007 and previously described MAS from mosquitoes collected in 2000-2005 for a portion of the S segment (nucleotides 144-605). Bootstrap values were assigned for 1000 replicates and are represented by the numbers on the branches. The previously described haplotypes are labeled. Colors represent the collection sites of the SI-mosquitoes (Purple: SVP/2006, Red: BEN2/2007, Green: NAT/2006 and NAT/2007, and Blue: CAL-GA/2007). The letters next to the MAS name distinguish different SI-mosquitoes collected from the same site.

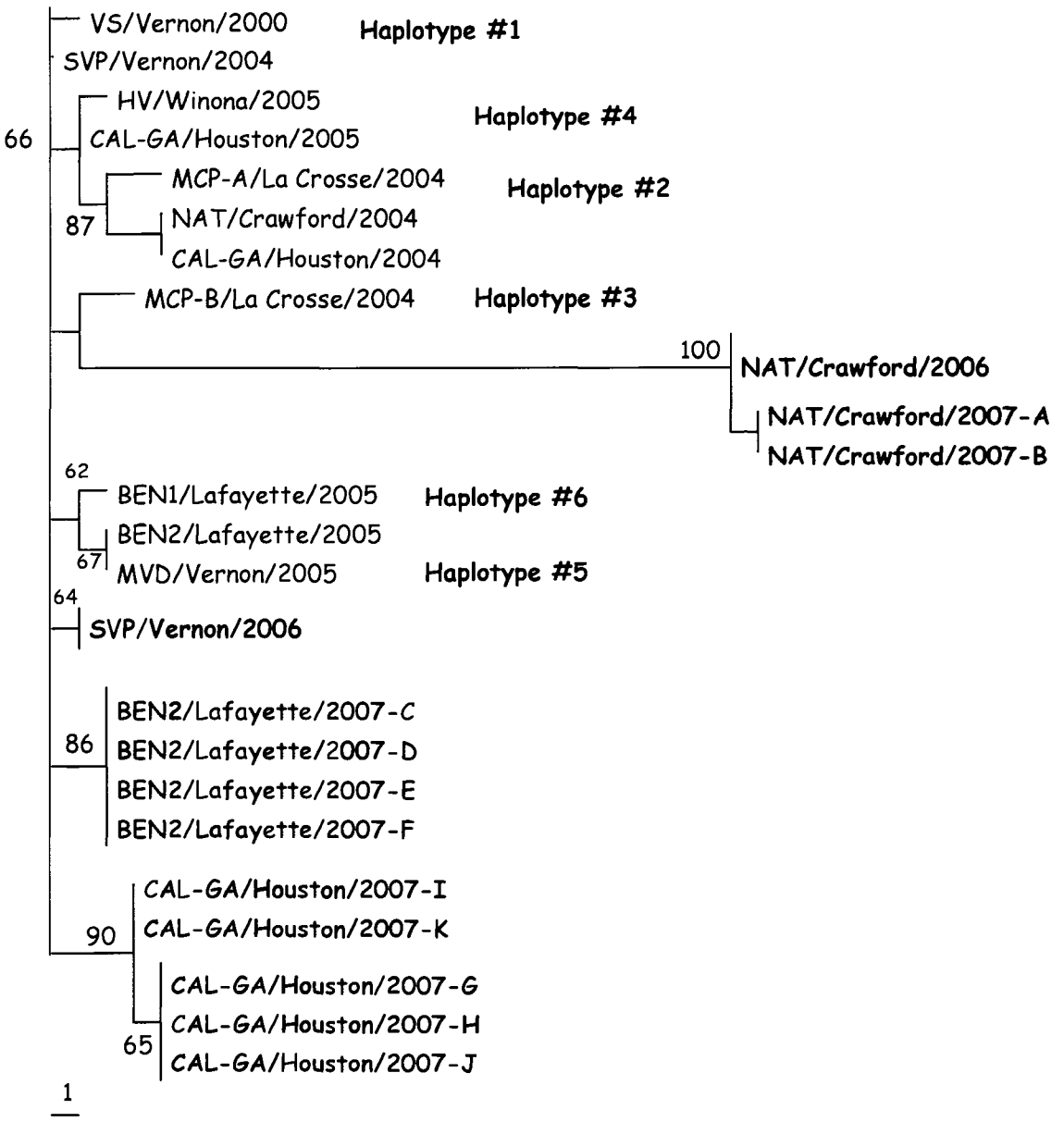



**Figure 4.8: Maximum parsimony phylogenetic analysis of a portion of the M segment RNA of MAS from normally-infected and super-infected field collected mosquitoes**

Phylogenetic analysis of MAS from field SI-mosquitoes collected in 2006 and 2007, and previously described MAS from mosquitoes collected in 2000-2005 for a portion of the M segment (nucleotides 1585-1995). Bootstrap values were assigned for 1000 replicates and are represented by the numbers on the branches. The previously described haplotypes are labeled. Colors represent the collection sites of the SI-mosquitoes (Purple: SVP/2006, Red: BEN2/2007, Green: NAT/2006 and NAT/2007, and Blue: CAL-GA/2007). The letters next to the MAS name distinguish different SI-mosquitoes collected from the same site.

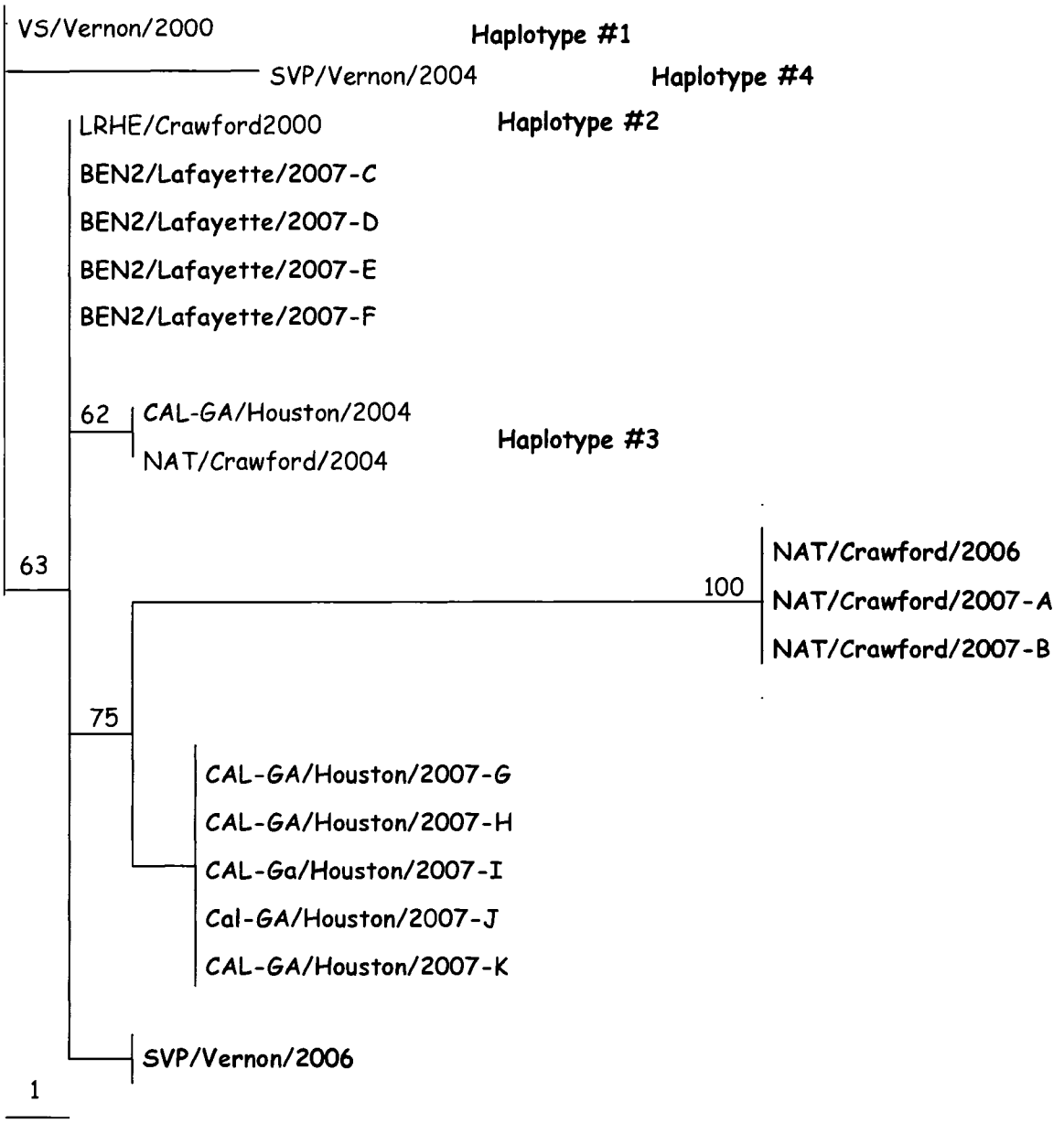



**Figure 4.9: Maximum parsimony phylogenetic analysis of a portion of the L segment RNA of MAS from normally-infected and super-infected field collected mosquitoes**

Phylogenetic analysis of MAS from field SI-mosquitoes collected in 2006 and 2007, and previously described MAS from mosquitoes collected in 2000-2005 for the L segment (nucleotide 140-626). Bootstrap values were assigned for 1000 replicates and are represented by the numbers on the branches. The previously described haplotypes are labeled. Colors represent the collection sites of the SI-mosquitoes (Purple: SVP/2006, Red: BEN2/2007, Green: NAT/2006 and NAT/2007, and Blue: CAL-GA/2007). The letters next to the MAS name distinguish different SI-mosquitoes collected from the same site.

analysis of the S, M and L segments, and MAS from mosquitoes collected in 2000-2005 (Table 4.3 and Figures 4.7-4.9). Through phylogenetic analyses of MAS obtained from mosquitoes collected in 2000-2005, distinct haplotypes were defined for S, M and L: 5 S haplotypes, 6 M haplotypes, and 4 L haplotypes (II. C.3., III. C.2.) (Figures 4.7-4.9).

The S segment maximum parsimony tree (Figure 4.7) has the distinct distribution of the defined haplotypes 1-5 (Table 4.3; II. C.3., III. C.2.). Interestingly, the MAS obtained from SI-mosquitoes collected from BEN2/2007 and SVP/2006 both cluster with haplotype #2, whereas the MAS obtained from mosquitoes from NAT/2006, NAT/2007 and CAL-GA/2007 form two new distinct clades. The tree morphology of the M segment (Figure 4.8) is quite different from the S segment. The defined haplotypes (Table 4.3; II., C.3., III., C.2.) form distinct clades; however the MAS obtained from the SI-mosquitoes do not associate in any of these clades and form four new clades.

The L segment maximum parsimony tree has similar tree topology observed in the M segment (Figure 4.9). The MAS from NAT/2006, NAT/2007, CAL-GA/2007 and SVP/2006 form new clades, whereas the MAS from BEN2/2007 cluster with haplotype #2.

**5. Sequence polymorphism, nucleotide substitution and nucleotide diversity**

Genome sequence polymorphism (theta) in a population is the proportion of nucleotide sites that are expected to be polymorphic in any sample in a particular region of the genome. The value of theta takes into account the number of nucleotide differences between each pair of sequences and the number of possible pairs (Kimura, 1969). The sequences of LACV isolates from the SI-mosquitoes were compared to published LACV isolates (Table 4.2). Genome sequence polymorphism for the S, M, and L segment

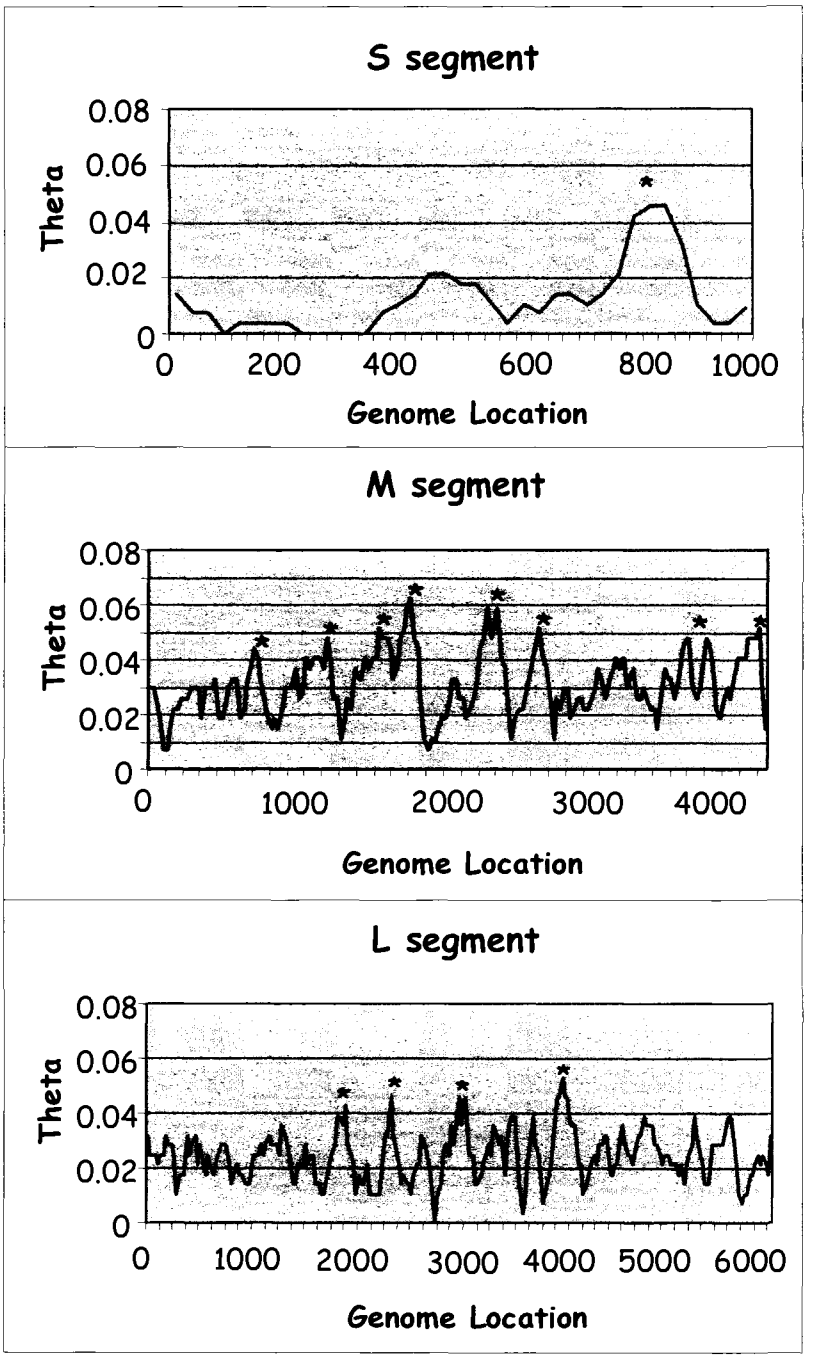



**Figure 4.10: Nucleotide polymorphism (theta) of the LACV S, M and L**

The asterisks represent the regions of significant nucleotide polymorphism. Polymorphic regions are regions of 100 nucleotides with 10 or greater polymorphisms.

sequences was 0.01282, 0.03218, and 0.02567 respectively. This can be seen graphically in Figure 4.10. In the S segment variability was found in nucleotides 730-900, which is mostly in the non-coding region (coding region is nucleotides 82-789). Variability in the M segment occurred in nucleotides 725-875, 1050-1300, 1475-2000, 2300-2600, 2700-3000, 3825-4000 and 4225-4475, which is mostly found in the coding region (coding region is nucleotides 62 – 4387). In the L segment, there were multiple regions of variability, nucleotides 1800-2000, 2300-2400, 3000-3200, and 4000-4250, which are all found in the coding region (nucleotides 62 – 6583).

The S, M and L segments had more synonymous substitutions than nonsynonymous, suggesting that the mutations are maintained by purifying selection (data not shown). The entire S segment had 13 synonymous substitutions and 1 nonsynonymous substitution. The entire M segment had 361 synonymous substitutions and 49 nonsynonymous substitutions. The entire L segment had 434 synonymous substitutions and 31 nonsynonymous substitutions.

**6. Nucleotide diversity**

The complete S, M and L sequences of the five virus isolates from the SI-mosquitoes were compared to previously published LACV sequences (Table 4.7-4.9). The most diverse sequences for the S segment are found in the isolates from NAT/2006 and NAT/2007 (NAT/Crawford, WI/Mosquito/2006 and NAT/Crawford, WI/Mosquito/2007), with 97.9% similarity with the isolates from Minnesota/Human/1960, Wisconsin/Mosquito/1977, Wisconsin/Human/1978-A, and Wisconsin/Human/1978-B (Table 4.7). The other isolates (BEN2/Lafayette, WI/Mosquito/2007, CAL-GA/Houston, MN/

Mosquito/2007 and SVP/Vernon, WI/Mosquito/2006) are 99.3% similar to previously isolated viruses.

Similar results are found with the M segment. The isolate, NAT/Crawford, WI/Mosquito/2007, is only 94.5% similar with the other SI isolates and previously isolated viruses (Table 4.8). The isolate, NAT/Crawford, WI/ Mosquito/2006 is not in the table due to incomplete sequence data. The isolates, SVP/Vernon, WI/ Mosquito/2006, CAL-GA/Houston, MN/Mosquito/2007 and BEN2/Lafayette, WI/Mosquito/2007, are about 99.3% similar to the previously published M segment sequences.

The nucleotide diversity analysis of L segment yielded similar results as the S and M segments (Table 4.9). The isolates from NAT/2006 and NAT/2007 have the most sequence diversity compared to the previously published sequences with 97.0% similarity. The isolates from SVP/2006, CAL-GA/2007 and BEN2/2007 are 99.2% similar to the isolates from Wisconsin and Minnesota. Interestingly, the isolate Wisconsin/ Mosquito/1977 is most similar to the isolates from NAT/2006 and NAT/2007, 99.2%, in the L segment whereas they are most similar to the isolates, Minnesota/Human/1960, Wisconsin/Human 1978-A and Wisconsin/Human/1978-B in the S and M segment. This suggests a possible reassortment event as seen with the maximum parsimony trees.

**7. Nucleotide sequence divergence**

The nucleotide sequence divergence between the virus isolates obtained from the SI-mosquitoes and the previously published virus isolates was determined (Figure 4.11). This analysis located regions of significant divergence between the two populations ($p < 0.10$). In the S segment, there are two regions of significant nucleotide divergence;

between nucleotides 1 and 25 and 959 and 984. Both of these regions are not in the coding regions of the nucleocapsid protein and the nonstructural protein. In the M segment, there are four regions of nucleotide divergence between the two populations; between nucleotides 900 and 925, nucleotides 1550 and 1700, nucleotides 2875 and 2900, and nucleotides 3500 and 3550. The first region of divergence is within the G2 glycoprotein and the other three are in the G1 glycoprotein. All of the polymorphisms in these regions are synonymous mutations. The L segment, there are four regions of divergence between the two populations; between nucleotides 1 and 25, nucleotides 2300 and 2600, nucleotides 3500 and 3600 and nucleotides 5600 and 5800. The first region is in the non-coding region and the other three regions are in the coding region for the RNA dependent RNA polymerase. All the polymorphisms in these regions were synonymous mutations.

## Table 4.7: Nucleotide diversity of the S segment comparing super-infected isolates and previously published sequences from the similar collection region

S segment

| | MN/Human/ 1960 | WI/Mosquito/ 1977 | WI/Human/ 1978-A | WI/Human/ 1978-B | BEN2/Lafayette, WI/Mosquito/ 2007 | SVP/Vernon, WI/Mosquito/ 2006 | CAL-GA/Houston, MN/Mosquito/2007 | NAT/Crawford, WI/Mosquito/ 2006 | NAT/Crawford, WI/Mosquito/ 2007 |
|---|---|---|---|---|---|---|---|---|---|
| Minnesota/Human/1960 | 0 | 1.00 | 0.997 | 0.995 | 0.993 | 0.994 | 0.988 | 0.979 | 0.979 |
| Wisconsin/Mosquito/1977 | | 0 | 0.997 | 0.995 | 0.993 | 0.994 | 0.988 | 0.979 | 0.979 |
| Wisconsin/Human/1978-A | | | 0 | 0.997 | 0.991 | 0.992 | 0.986 | 0.977 | 0.977 |
| Wisconsin/HUman/1978-B | | | | 0 | 0.993 | 0.994 | 0.988 | 0.979 | 0.979 |
| BEN2/Lafayette, WI/Mosquito/2007 | | | | | 0 | 0.996 | 0.990 | 0.983 | 0.983 |
| SVP/Vernon, WI/Mosquito/2006 | | | | | | 0 | 0.991 | 0.982 | 0.982 |
| CAL-GA/Houston, MN/Mosquito/2007 | | | | | | | 0 | 0.980 | 0.980 |
| NAT/Crawford, WI/Mosquito/2006 | | | | | | | | 0 | 1.00 |
| NAT/Crawford, WI/Mosquito/2007 | | | | | | | | | 0 |

## Table 4.8: Nucleotide diversity of the M segment comparing super-infected isolates and previously published sequences from the similar collection region

M segment

| | MN/Human/ 1960 | WI/Mosquito/ 1977 | WI/Human/ 1978-A | WI/Human/ 1978-B | BEN2/Lafayette, WI/Mosquito/ 2007 | SVP/Vernon, WI/Mosquito/ 2006 | CAL-GA/Houston, MN/Mosquito/2007 | NAT/Crawford, WI/Mosquito/ 2007 |
|---|---|---|---|---|---|---|---|---|
| Minnesota/Human/1960 | 0 | 0.998 | 0.996 | 0.995 | 0.993 | 0.994 | 0.993 | 0.946 |
| Wisconsin/Mosquito/1977 | | 0 | 0.995 | 0.994 | 0.992 | 0.993 | 0.992 | 0.945 |
| Wisconsin/Human/1978-A | | | 0 | 0.999 | 0.994 | 0.995 | 0.993 | 0.944 |
| Wisconsin/HUman/1978-B | | | | 0 | 0.994 | 0.994 | 0.993 | 0.944 |
| BEN2/Lafayette, WI/Mosquito/2007 | | | | | 0 | 0.993 | 0.991 | 0.943 |
| SVP/Vernon, WI/Mosquito/2006 | | | | | | 0 | 0.992 | 0.943 |
| CAL-GA/Houston, MN/Mosquito/2007 | | | | | | | 0 | 0.943 |
| NAT/Crawford, WI/Mosquito/2007 | | | | | | | | 0 |

## Table 4.9: Nucleotide diversity of the L segment comparing super-infected isolates and previously published sequences from the similar collection region

L segment

| | MN/Human/ 1960 | WI/Mosquito/ 1977 | WI/Human/ 1978-A | WI/Human/ 1978-B | BEN2/Lafayette, WI/Mosquito/ 2007 | SVP/Vernon, WI/Mosquito/ 2006 | CAL-GA/Houston, MN/Mosquito/2007 | NAT/Crawford, WI/Mosquito/ 2006 | NAT/Crawford, WI/Mosquito/ 2007 |
|---|---|---|---|---|---|---|---|---|---|
| Minnesota/Human/1960 | 0 | 0.971 | 0.994 | 0.993 | 0.994 | 0.993 | 0.993 | 0.946 | 0.971 |
| Wisconsin/Mosquito/1977 | | 0 | 0.969 | 0.968 | 0.970 | 0.969 | 0.968 | 0.967 | 0.992 |
| Wisconsin/Human/1978-A | | | 0 | 0.999 | 0.992 | 0.991 | 0.991 | 0.944 | 0.969 |
| Wisconsin/HUman/1978-B | | | | 0 | 0.992 | 0.991 | 0.991 | 0.945 | 0.969 |
| BEN2/Lafayette, WI/Mosquito/2007 | | | | | 0 | 0.991 | 0.992 | 0.945 | 0.970 |
| SVP/Vernon, WI/Mosquito/2006 | | | | | | 0 | 0.991 | 0.945 | 0.970 |
| CAL-GA/Houston, MN/Mosquito/2007 | | | | | | | 0 | 0.944 | 0.969 |
| NAT/Crawford, WI/Mosquito/2006 | | | | | | | | 0 | 0.972 |
| NAT/Crawford, WI/Mosquito/2007 | | | | | | | | | 0 |

The entire S, M and L segment sequences isolated from the SI-mosquitoes were aligned with the published virus sequences obtained from LACV isolates recovered from Wisconsin and Minnesota. The M segment sequence from NAT/Crawford, WI/Mosquito/2006 was not included in this analysis because the entire segment was not sequenced.

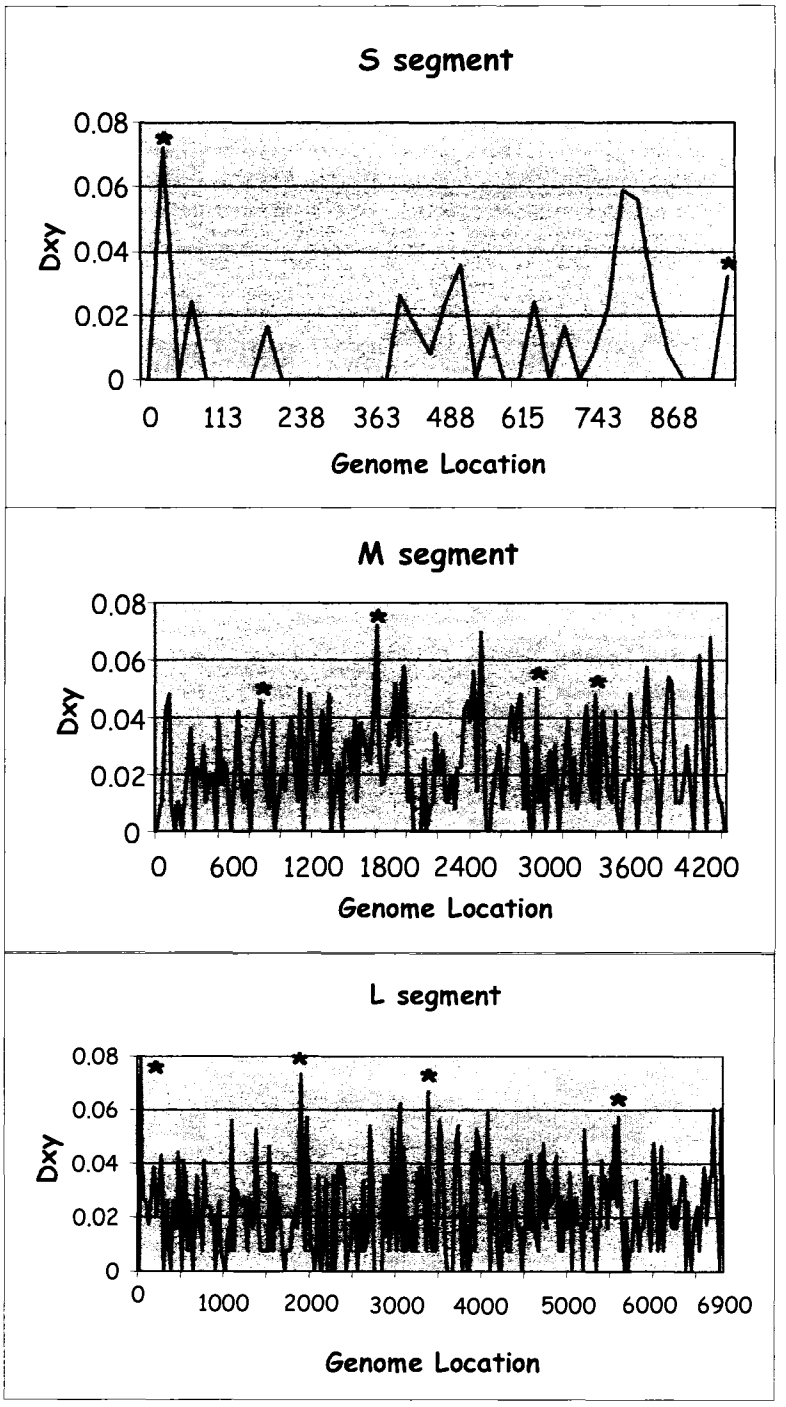



**Figure 4.11: Nucleotide sequence divergence ($D_{xy}$) of the LACV S, M and L**

The asterisks represent the regions of significant nucleotide divergence between the two populations. Significant nucleotide divergence regions are regions of 25 nucleotides consisting of polymorphisms in one population but not the other.

## D. DISCUSSION

The identification of super-infected LACV mosquitoes in the field suggests that LACV could be maintained in the population through a stabilized infection. This is an exciting possibility because it could explain how LACV can persist in nature with low field infection rates. These SI-mosquitoes were detected in the population at a low level (0.08%) (Table 4.2). However they were widely distributed in the collection area (Figure 4.1), which suggests that this is not a localized phenomenon within one collection site. SI-mosquitoes were collected from SVP/2006 (Vernon County, WI), CAL-GA/2007 (Houston County, MN) and BEN2/2007 (Lafayette County, WI) and NAT/2006 and NAT/2007 (Crawford County, WI) (Figure 4.1). The identification of a SI-mosquito in the same site in two different years suggests this genotype of LACV is stabilized in mosquitoes in this collection site. The prevalence of SI-mosquitoes ranged from 0.84% (NAT/2007) to 12.0% (CAL-GA/2007) (Table 4.5). If these SI-mosquitoes transmit LACV to all or most of their progeny, it is conceivable that the virus could be maintained in nature at a fairly constant level by a relatively small number of stabilized females (Tesh & Shroyer, 1980).

Phylogenetic studies were then conducted to characterize the viruses isolated from SI-mosquitoes and to investigate the genetic relatedness of the LACV isolated from SI-mosquitoes to previously isolated viruses (Table 4.2). A maximum parsimony tree was created using the entire sequence of the S, M and L segments (Figures 4.4-4.6).

The maximum parsimony analyses of the S and L segments contain previously published LACV sequences and sequences from the viruses isolated from the SI-mosquitoes (Table 4.3) (Figures 4.4 and 4.6). In the S and L trees, the virus isolates from

the SI-mosquitoes collected from NAT/2006 and NAT/2007 form their own clade. Interestingly this clade also contains the isolate North Carolina/Mosquito/1977 with strong bootstrap support. This is suggestive of a diverse LACV lineage. Previous evidence has suggested that there is isolation by distance regionally. There are lineages containing viruses from WI and MN and lineages containing viruses from other regions of the United States and they do not typically cluster together (Armstrong & Andreadis, 2006, Klimas et al., 1981). However, this phylogenetic analysis yielded different results that suggest there is no isolation by distance regionally. This was further emphasized in the M segment maximum parsimony analysis because more sequences were available (Table 4.2, Figure 4.5) The M segment sequences clearly distributed into four different clades. Clade #2 contained viruses recovered from Tennessee and North Carolina and clade #4 contained the viruses recovered from New York, Alabama, Connecticut and Georgia. Clades #1 and #3 contained most of the isolates from WI and MN. However, the isolate from Missouri clustered into clade #1 and the isolates from West Virginia, Ohio and North Carolina clustered into clade #3 with the isolates from NAT/2006 and NAT/2007 and one isolate from Wisconsin. These results differ from previous findings suggesting isolation by distance on a regional scale. These results suggest that while there is significant variation observed between all of the isolates, isolates from different regions of the country do cluster together. The reasons for this are unknown. A wide scale analysis with multiple isolates from multiple regions of the country needs to be performed to truly understand the genetic determinants of geographic distance on a regional scale.

The possibility of segment reassortment was also observed in the phylogenetic analyses (Figures 4.4 and 4.6). The isolate Wisconsin/Mosquito/1977 clustered with Minnesota/Human/1960, Wisconsin/Human/ 1978-A, Wisconsin/Human/ 1978-B, Wisconsin/Mosquito/1977 and the isolates from BEN/2007, SVP/2006, and CAL-GA/2007 in the S segment tree. However, in the L segment tree, Wisconsin/Mosquito/ 1977 clustered with the isolates from NAT/2006 and NAT/2007, which suggests possible segment reassortment. This was also observed in the nucleotide diversity analysis (Table 4.7-4.9). In the S segment nucleotide diversity analysis, Wisconsin/Mosquito/1977 is most similar (99.7-100% similarity) to the isolates: Minnesota/Human/1960, Wisconsin/Human/1978-A, Wisconsin/Human/ 1978-B, Wisconsin/Mosquito/1977 and the isolates from BEN2/2007, SVP/2006, and CAL-GA/2007. The L segment nucleotide diversity analysis revealed that Wisconsin /Mosquito/1977 is 99.2% similar to the isolates from NAT/2006 and NAT/2007 and only 97% similar to the other isolates, suggesting possible reassortment. As discussed in II.C.5., there is significant genetic variation of LACV due to segment reassortment. This further supports the hypothesis that LACV is highly variable in nature and is constantly evolving through both genetic drift and segment reassortment.

It is provocative that in the M segment phylogenetic analysis a virus isolate from a SI-mosquito and an isolate from a known human fatality clustered together. The BEN2/Lafayette, WI/ Mosquito/2007 and Missouri/Human/1993 isolates are genetically similar as evidenced by the location on the tree (Figure 4.5). Perhaps the viruses causing super infections in mosquitoes may cause severe clinical outcomes in humans. This could be due to dose of inoculum in the saliva since the average titer of the SI-mosquitoes is

extremely high (Table 4.6) or to linkage of virus replication and persistence in mosquitoes and efficiency with human replication.

In the S, M, and L segment maximum parsimony analysis, isolates from the SI-mosquitoes collected from NAT/2006 and NAT/2007 (NAT/Crawford, WI/Mosquito 2006, NAT/Crawford, WI/ Mosquito/2007) did not cluster with other isolates from the SI-virus mosquitoes (Figure 4.5). The isolates from NAT/2006 and NAT/2007 are clearly a different genotype than the other viruses isolated from SI-mosquitoes.

It is also noteworthy that the viruses isolated from SI-mosquitoes collected from NAT in 2006 and 2007 (NAT/Crawford, WI/Mosquito 2006, NAT/Crawford, WI/ Mosquito/2007) have similar S, M and L segment sequences (Figures 4.4-4.6). This conflicts with previous data presented in III. (C.6. and C.7.), suggesting that the haplotypes may evolve each year. One possible reason for this could a stabilized infection. The NAT/2006 and NAT/2007 isolates may reflect a local mosquito lineage. The stabilized viral haplotype could be maintained in the population each year possibly through purifying or stabilizing selection in which the genetic diversity decreases as the virus stabilizes in the population.

A maximum parsimony analysis was also performed for the S, M and L segments with MAS obtained from the SI-mosquitoes (instead of virus isolates) and with MAS obtained from mosquitoes collected in 2000-2005 (Figure 4.7-4.9) (previously described in II.C.3. Figure 2.2-2.4 and III. C.2., Figure 3.2). The maximum parsimony analyses of the S, M and L segments of the MAS yielded similar results found in II.C.4, and II.C.5., suggesting evolution through segment reassortment or genetic drift. In the L segment maximum parsimony analysis (Figure 4.9), the MAS obtained from the SI-mosquitoes

collected from NAT/2006, NAT/2007, SVP/2006 and CAL-GA/2007, all form new clades, whereas the MAS from SI-mosquitoes from BEN2/2007 cluster in clade #2. The M segment maximum parsimony analysis reveals all four MAS from SI-mosquitoes cluster into separate, genetically diverse clades (Figure 4.8). In the S segment analysis, the MAS from SI-mosquitoes from SVP/2006 cluster into haplotype #2 (Figure 4.7) and cluster into previously undefined clades in the M and L segment analyses (Figures 4.8 and 4.9). The MAS obtained from SI-mosquitoes collected from BEN2/2007 cluster with haplotype #2 (Figure 4.7) and forms a new clade in the M segment tree (Figure 4.8). In all the S, M and L trees, the MAS obtained from mosquitoes collected from NAT/2006, NAT/2007 and CAL-GA/2007 cluster into individual, diverse clades (Figures 4.7-4.9). The genetic diversity observed with the SI-mosquitoes in the MAS phylogenetic analyses could be due to a few factors. As demonstrated in III.C.6.a., III.C.6.b., and III.C.7., there is a significant association with collection year and LACV sequence. The variability could be due to different collection years. It could also be due to viral determinants of super infection. The variability observed between the S, M and L segment trees could also be indicative of genetic drift or segment reassortment, resulting in the establishment of a new clade.

The three maximum parsimony tree analyses reveal that the MAS obtained from the SI-mosquitoes do not cluster together with MAS obtained from other mosquitoes in the same collection sites. For example, in the S segment tree (Figure 4.7), the MAS from BEN2 in 2007 is found in haplotype #2, whereas the MAS obtained from BEN2 in 2005, is found in haplotype #5. This suggests that either there are multiple genotypes present in

one collection site or the haplotypes are evolving each year as observed in III.C.6.a., III.C.6.b., and III.C.7.

The nucleotide divergence analysis of the S, M and L segments reveals possible regions responsible for super infection. The S segment contained two regions of nucleotide divergence (Figure 4.11), which are both in the non-coding region. It remains unknown whether these regions contribute to infection, however they may condition replication and transmission efficiency. The M segment had four regions of significant divergence, one region in the G2 glycoprotein and three regions in the G1 glycoprotein. The polymorphisms contributing to nucleotide divergence were synonymous mutations, which indicate they may not influence super infection. The L segment also had four regions of nucleotide divergence, three of which were in the coding region for the RNA dependent RNA polymerase. As with the M segment, all the polymorphisms were synonymous. This particular analysis did not indicate any regions of significant nucleotide divergence resulting in nonsynonymous mutations. However, the sample size was small and therefore some regions may be contributing to super infection, but the analysis was not sensitive enough to detect them. Further analyses need to determine if there are genetic determinants of LACV contributing to super infection.

The S segment phylogenetic analyses of both the virus isolates and the MAS could suggest a possible role in super infection (Figures 4.4 and 4.7). In both these phylogenetic trees, the isolates and MAS from the SI-mosquitoes cluster into the same lineage. In the M and the L phylogenetic analyses, the isolates and MAS cluster into different lineages. This could be indicative of a possible role the S segment may play in super infection.

In order to truly understand the genetic determinants of super infection, the LACV isolated from the SI-mosquitoes needs to be further analyzed. The M and L segments play the primary role in virulence and replication and are possibly contributing to super infection. The M segment, which codes for two glycoproteins, G1 and G2 and a nonstructural protein, is responsible for virulence, host range, tissue tropism, transmissibility, neutralization, hemagglutination, and membrane fusion (Beaty et al., 1982, Gonzalez-Scarano et al., 1988). The L segment codes for the RNA dependent RNA polymerase and is the major determinant of replication efficiency (Endres et al., 1989). Studies have also suggested that virulence of LACV is under polygenic control (Janssen et al., 1986). The M segment is the major determinant of mouse virulence and mosquito infectivity, but the S and L segments could modulate the dominant effect of the M segment of a nonneuroinvasive California serogroup virus. Determination of conserved regions in the S, M and L segments of the SI virus isolates that differ from viruses isolated from NI-mosquitoes could reveal possible viral genetic determinants of super infection, which could be manipulated and analyzed through the use of a reverse genetics system (Blakqori & Weber, 2005).

In this study, isolation of LACV from NI-mosquitoes was not successful. This could be due to a variety of reasons. Eggs were collected in the field and stored in a hot warehouse until shipped, over the course of a few days to Colorado. As soon as the eggs reached AIDL, they were placed in the insectary, hatched and reared. Environmental factors in the collection and shipping process could contribute to loss of virus titer. The titer of the SI-mosquitoes (Table 4.6) is approximately 3.78 $\log_{10}TCID_{50}$/ml. This would be lower in NI-mosquitoes. An additional complication could have been the assay. In

previous studies, virus was isolated by inoculation of samples into suckling mouse brains. Cell culture assays are likely not as sensitive. Low virus titer, titer loss during processing, and insensitive assays, undoubtedly contributed to the inability to isolate virus NI-mosquitoes. Future studies need to process the mosquito eggs quickly and use cold chain in order to assure virus isolation from NI-mosquitoes.

Additionally, LACV mating studies with the newly discovered SI-mosquitoes need to be performed to determine the filial infection rates. If the SI-mosquitoes transmit LACV to most or all of their progeny, they may represent a stabilized infection. This, however, will be difficult to accomplish due to a variety of factors. The low percentage of the SI-mosquitoes present in the field is problematic. Once the SI-mosquitoes are identified, they are difficult to maintain and colonize in a laboratory setting and mating success is low. Field collected mosquitoes do not mate or feed on artificial blood meals as well as laboratory colonized females (Mather & Defoliart, 1984). However, such studies are crucial to determine if SI-mosquitoes and their progeny have a filial infection rate that can maintain LACV through a stabilized infection.

Stabilized infection was first observed with Sigma virus and *Drosophila melanogaster* (Seecof, 1968). When the virus was vertically transmitted to a small percentage of the progeny by infecting the developing oocytes directly, the virus was in a "nonstabilized" state (Fleuriet, 1988). However, if the germarium of a female became infected, she transmitted the virus to nearly 100% of her progeny, which was designated as a "stabilized" state. It would be provocative if this was occurring with *Ae. triseriatus* and leading to super infection. Additional experiments are needed to determine if the germarium is indeed infected, resulting in majority of the progeny infected.

Mathematical models of LACV and KEY have strongly suggested that LACV cannot persist in nature (Defoliart, 1983, Fine & LeDuc, 1978) due to low field and filial infection rates (Beaty & Thompson, 1975, Lisitza et al., 1977, Watts et al., 1974). However, these models have not considered such factors as stabilized infection or mating advantages (discussed in V.). These two factors may help maintain stable LACV prevalence from year to year in the vector population.

This study suggests that LACV could be maintained in nature through stabilized infection in a small number of females. It is not known whether the super infection is due to the mosquito genetics, viral genetics or possibly the interaction between the mosquito and virus. Further studies are needed to begin to understand this unique relationship. Specific mosquito genes, metabolic pathways, or immune pathways could also condition for LACV super infection. Other experiments such as comparison of gene expression using microarrays, genetic crossing experiments, mating experiments to determine filial infection rates and the pathogenesis of LACV in super-infected mosquitoes also need to be performed. Further studies are needed to elucidate the viral and vector genetic determinants of super infection and to determine the potential epidemiological significance of this phenomenon.

# V. COMPARISON OF INSEMINATION RATES OF LA CROSSE VIRUS TRANSOVARIALLY INFECTED AND UNINFECTED *AEDES TRISERIATUS* MOSQUITOES

## A. INTRODUCTION

La Crosse virus (LACV) is maintained in a cycle between *Aedes triseriatus* mosquitoes (Eastern tree hole mosquito) and chipmunks and tree squirrels (Calisher, 1994, Watts et al., 1972). The virus can also be transmitted between mosquitoes horizontally (venereally) and vertically (transovarially), thereby enabling LACV to overwinter in the eggs of an infected mosquito (Thompson & Beaty, 1978, Watts et al., 1973, Watts et al., 1974) Transovarial transmission (TOT) is an extraordinary method of amplification since an infected female will produce many infected progeny in a lifetime. The newly hatched, infected females are capable of transmitting LACV upon emergence. Infection through TOT could have a significant effect on viral amplification and maintenance because any behavioral or physiological changes due to LACV infection will be evident as soon as the mosquito emerges. Mosquito behavior modification resulting from LACV infection could influence viral transmission and possibly amplify the viral prevalence in the environment.

There is strong evidence of behavioral changes in mosquitoes due to viral infection. Pathogens may induce many mosquito adaptations that are beneficial for survival and dispersal. For example, LACV-infected *Ae. triseriatus* females probe more and engorge less than uninfected mosquitoes (Grimstad et al., 1980). In one study, 21% of infected females probed once to obtain a partial bloodmeal and 79% probed multiple times to obtain a partial bloodmeal. In the uninfected females, 52% were fully engorged following one probe and 48% were fully engorged following multiple probes (Grimstad et al., 1980). Increased probing and partial engorgement could result in increased viral transmission and amplification.

A similar association was found with dengue-infected *Ae. aegypti* mosquitoes (Platt et al., 1997). The total time required for feeding and the mean time spent probing are both significantly longer in infected mosquitoes than in uninfected ones. This increased time required by infected mosquitoes to acquire a bloodmeal may contribute to the efficiency of *Ae. aegypti* as a vector of dengue virus. In addition, longer feeding periods are more likely to be interrupted by the host, which increases the chance that an infected mosquito will probe or feed on additional hosts (Platt et al., 1997).

The flight activity of mosquitoes has also been altered through infection with pathogens. *Ae. aegypti* mosquitoes infected with *Dirofilaria immitis* have an altered flight activity compared to uninfected mosquitoes (Berry et al., 1987). Mosquitoes infected with more than four larvae become more active than uninfected mosquitoes eight days after infection possibly due to irritation related to the activity of the larvae. Flight activity of mosquitoes infected with more than four larvae is suppressed on days 10 and 14 post-infection, corresponding to times of greatest disruption of the Malpighian tubules by the developing larvae. Decreased flight activity also occurs in *Culex tarsalis* when infected with Western Equine encephalitis virus (Lee et al., 2000).

LACV infection of *Ae. triseriatus* mosquitoes also produces behavioral changes that could promote virus amplification and maintenance in nature. An increased insemination rate was observed in female *Ae. triseriatus* mosquitoes that were orally infected with LACV in the laboratory (Gabitzsch et al., 2006). Mosquitoes were given an infectious bloodmeal followed by a non-infectious bloodmeal 14 days later and then allowed to mate. Of the mosquitoes ingesting the second bloodmeal, 91.9% and 78.0% (trials 1 and 2) of LACV-infected mosquitoes were inseminated compared to 71.7% and

30.0% (trials 1 and 2) of uninfected mosquitoes. Similar results were observed for the mosquitoes that did not ingest the second bloodmeal; 91.7% and 56.8% (trials 1 and 2) of the LACV-infected mosquitoes were inseminated compared to 50.0% and 27.3% (trials 1 and 2) insemination rates for the uninfected mosquitoes.

Multiple factors may contribute to the increased rate of insemination. Differences in viral titers present within a mosquito can affect transmission and possibly behavior (Watts et al., 1972). Virus infection could also affect the pheromones of the infected female mosquitoes. Differences in certain pheromones between females lead to differences in mating efficiency (Anthony & Jallon, 1982, Ferveur et al., 1996, Jallon, 1984). Additionally, the female accessory gland could contribute to increased mating efficiency. The single accessory gland is a globular structure situated dorsolaterally to the rectum and posterior to the spermathecae. Its duct runs to the spermathecal vestibule, where it opens into the upper vagina at a point just above the orifice of the spermathecal ducts (Clements, 1992). The accessory gland is composed of secretory cells, each of which surrounds a large extracellular space filled with secretion. The extracellular space is connected to the accessory gland duct by a cuticular canal. The role of the accessory gland in mating is currently unknown, but if the gland becomes infected during a disseminated infection, it may somehow condition increased mating efficiency.

Based on previous evidence demonstrating increased insemination rates in mosquitoes orally-infected with LACV (Gabitzsch et al., 2006), further studies were conducted to test the hypothesis that transovarial infection by LACV is associated with increased mating efficiency in *Ae. triseriatus* females. This could amplify LACV in nature; i.e. progeny of LACV-infected mosquitoes would have a selective advantage over

those of uninfected females. Possible determinants for this association were also investigated.

## B. METHODS AND MATERIALS

### 1. Laboratory mosquitoes

*Aedes triseriatus* mosquitoes (AIDL colony) were originally collected as eggs near La Crosse, WI in 1981 and have been continuously maintained in the Arthropod-borne and Infectious Diseases Laboratory at Colorado State University (Fort Collins, CO). Mosquitoes that have been selected for increased rates of LACV TOT and have a TOT rate of approximately 50% (L+T+ AIDL colony mosquitoes) were also used in these experiments (Graham et al., 1999). All eggs were hatched and larvae were fed a 1:1 mixture of ground fish food:mouse food. Pupae were separated by sex and allowed to emerge in 3.8 liter containers. Adults were maintained on sugar cubes and water at 20-23°C, 80% relative humidity and a photoperiod of 16:8 (L:D) hours.

### 2. Field mosquitoes

*Ae. triseriatus* eggs were collected from oviposition traps at various collection sites in Wisconsin and Minnesota in the summer of 2006 (III. B.1., Figure 3.1). The eggs were transported to the insectaries at AIDL and promptly hatched and reared. Female pupae were allowed to emerge in 3.8 liter containers. Male mosquitoes were not used in these experiments.

### 3. Immunofluorescence assay

Mosquito legs were removed, squashed onto microscope slides, fixed in acetone and assayed for LACV antigen by direct immunofluorescence assay (IFA) using a LACV-specific polyclonal antibody at a dilution of 1:100 (Beaty & Thompson, 1975).

**4. Bloodmeals**

In triplicate trials, the AIDL and L+T+ mosquitoes were separately given a bloodmeal containing defibrinated sheep blood via a membrane feeder (Rutledge et al., 1964). The mosquitoes were allowed to feed for 1.5h, after which the unengorged females were removed. The field collected mosquitoes were given a bloodmeal and the engorged females were then tested for LACV antigen by IFA of leg tissue.

**5. Mosquito mating**

Females from eight different groups [(1) AIDL, no bloodmeal, (2) L+T+, no bloodmeal, (3) AIDL, bloodmeal, (4) L+T+, bloodmeal, (5) Field LACV-, no bloodmeal, (6) Field LACV+, no bloodmeal, (7) Field LACV-, bloodmeal, and (8) Field LACV+, bloodmeal] (Table 5.1) were placed separately into 0.27-$m^3$ cages. Laboratory colonized male mosquitoes were added to each cage of laboratory colonized and field collected female mosquitoes at a 2:1 (male:female) ratio, which is optimal for insemination (Mather & Defoliart, 1984). Laboratory female mosquitoes were removed from the cage every 24 hours beginning at one day post-mixing. The field collected females were removed seven days post-mixing to determine insemination rates. The spermathecae of all females were examined microscopically for the presence of motile sperm to determine if insemination had occurred (Gabitzsch et al., 2006, Mather & Defoliart, 1984). The experiments were carried out for seven days.

**Table 5.1: Mosquitoes and conditions used for insemination and cuticular hydrocarbon experiments**

| | |
|---|---|
| AIDL colony LAC- | No |
| AIDL colony LAC- | Yes |
| L+T+ LAC+ | No |
| L+T+ LAC+ | Yes |
| Field-collected LAC- | No |
| Field-collected LAC- | Yes |
| Field-collected LAC+ | No |
| Field-collected LAC+ | Yes |

**6. Determination of genome equivalents through quantitative reverse transcription PCR (qRT-PCR)**

**a. RNA isolation and cDNA synthesis**

Mosquito carcasses were triturated in 500µl of Trizol (Invitrogen) using a pellet pestle (Fisher Scientific) and placed at -70°C for >1 hour. Samples were thawed to room temperature and RNA was extracted according to manufacturer's instructions. Approximately 100ng of total RNA was mixed with a reverse transcription primer specific for the LACV S segment (5'-TCA AGA GTG TGA TGT CGG ATT TGG-3') (Kempf et al., 2006). The primer binds to nucleotides 71-95 of LACV S segment mRNA. Superscript II (Invitrogen) was used to produce a cDNA containing the reverse transcript, which has the sequence of LACV genomic RNA. In brief, RNA, 10µM dNTPs and 50µM primer were incubated at 65°C for 10 minutes. Superscript II, 5x first-strand buffer, and 1mM DTT were added and incubated at 42°C for 50 minutes. The reaction was deactivated at 70°C for 15 minutes.

**b. Q-PCR analysis**

One-fourth of the cDNA prepared from each reverse transcription reaction was used for Q-PCR analysis. The forward primer was SF (5'-GGT TAG CCT TCC TCT

CTG GCT TA-3') which binds to nucleotides 246-268 of the LACV S mRNA. The reverse primer, SR (5'- CCT TGC TGC AGT TAG GAT CTT CTT-3') binds to nucleotides 186-209 of the LACV S mRNA (Kempf et al., 2006). Q-PCR primers and probes were purchased from Qiagen and Q-PCR reagents were obtained from Stratagene (Brilliant Q-PCR reagents with SureStart *Taq* DNA polymerase) and used according to manufacturer's instructions (with the exception of using 20µl reactions instead of 50µl reactions). Serial 10-fold dilutions of known copy number control plasmid ($1 \times 10^1$ to $1 \times 10^8$) were amplified simultaneously to generate standard curves.

**7. Cuticular hydrocarbon analysis**

The cuticular hydrocarbon (CHC) profiles were compared for four different groups: (1) non-bloodfed AIDL colony mosquitoes, (2) non-bloodfed L+T+ mosquitoes, (3) bloodfed AIDL colony mosquitoes and (4) bloodfed L+T+ mosquitoes (Table 5.1). CHCs from pools of 10 mosquitoes were analyzed at various time points either post-bloodmeal or post-eclosion (0, 24, 48 and 72 hours). The methods used were adapted from Polerstock *et al.* (Polerstock et al., 2002). CHCs were extracted in 200µl of hexane for 5 minutes in 15-ml conical glass tubes. This was repeated twice. The extract was filtered in glass wool to remove contaminants and then concentrated under a stream of nitrogen. Immediately before their injection into the gas chromatography and mass spectrophotometer (GC/MS), the concentrated extracts were reconstituted in 15µl of hexane. All samples were analyzed within 1 hour of extraction. One µl was analyzed in a Trace GC 2000 gas chromatograph (Finnigan-Thermoquest, San Jose, CA) fitted with a DB-5 column (10 m x 0.18 mm, Agilent Technologies, Palo Alto, CA) linked to a Polaris Mass Detector (Finnigan-Thermoquest). Analytical runs were programmed at an initial

temperature of 80°C, held for 1 minute, which was first raised to 130°C at a rate of 30°C per minute followed by at a rate of 10°C per minute to a final temperature of 280°C and held for 10 min.

**8. Female accessory sex gland dissection and staining**

Female *Ae. triseriatus* mosquitoes were administered an infectious bloodmeal (7.0 $\log_{10}TCID_{50}$/ml). Fourteen days post-bloodmeal the females were anesthetized and the female accessory sex gland was removed (Figure 5.1). The presence of LACV antigen in the accessory sex gland was determined through immunofluorescence assay (described above).

**Figure 5.1: Female accessory sex gland of *Ae. Triseriatus* (100x)**

**9. Statistics**

A student's t-test was used to compare the mean LACV genome equivalents in field collected inseminated and non-inseminated mosquitoes. This test is typically used to determine if the means of two normally distributed populations differ significantly. The Wilcoxon signed rank test was used to determine if there was a significant difference in genome equivalents between the L+T+ inseminated and non-inseminated mosquitoes. A

Wilcoxon test is used when the population is not normally distributed and this is the non-parametric alternative to the Student's t-test (Wilcoxon, 1945).

## C. RESULTS

### 1. Comparison of insemination rates of laboratory colonized non-bloodfed LACV- and LACV+ mosquitoes

In the non-bloodfed mosquitoes, the LACV-TOT+ laboratory colonized females were inseminated more rapidly than the uninfected mosquitoes. Motile sperm was infrequently detected in the spermathecae of uninfected and infected mosquitoes on days 1 and 2 (data not shown).

There was a significant difference in insemination rate between LACV- and LACV+ by day 4 when the LACV+ mosquitoes had an insemination rate of 49.5% (n = 200) compared to 33.9% (n = 280) for the uninfected mosquitoes ($\chi 2 = 11.75$, $p < 0.001$) (Microsoft Excel) (Figure 5.2). By day 7, the rate of insemination was similar for both the LACV+ (83.2%, n = 191) and LACV- (80.6%, n = 268) mosquitoes ($\chi 2 = 0.523$, $p = 0.469$).

### 2. Comparison of insemination rates of laboratory colonized bloodfed LACV- and LACV+ mosquitoes

In bloodfed mosquitoes, the LACV+ females were inseminated more rapidly than LACV- females. The rate of insemination was minimal in the first day post-mixing (data not shown). There was a significant difference in insemination rates observed on day 3 (Figure 5.3), when LACV+ mosquitoes had an insemination rate of 68.2% (n = 170) compared to 39.6% of LACV- (n = 192) ($\chi 2 = 29.7$, $p = <0.001$). By day 6, the

insemination rates for both the infected and uninfected groups had equalized to 85.3% (n = 199) and 88.4% (n = 258), respectively ($\chi 2 = 0.977$, $p = 0.323$) (Figure 5.3).

**3. Comparison of insemination rates of non-bloodfed LACV- and LACV+ field collected mosquitoes**

In the non-bloodfed field collected mosquitoes, the LACV+ females were inseminated more often then the LACV- females. These field collected mosquitoes were allowed to mate for seven days because insemination rates are much lower compared to the laboratory colony mosquitoes. The LACV+ mosquitoes had an insemination rate of

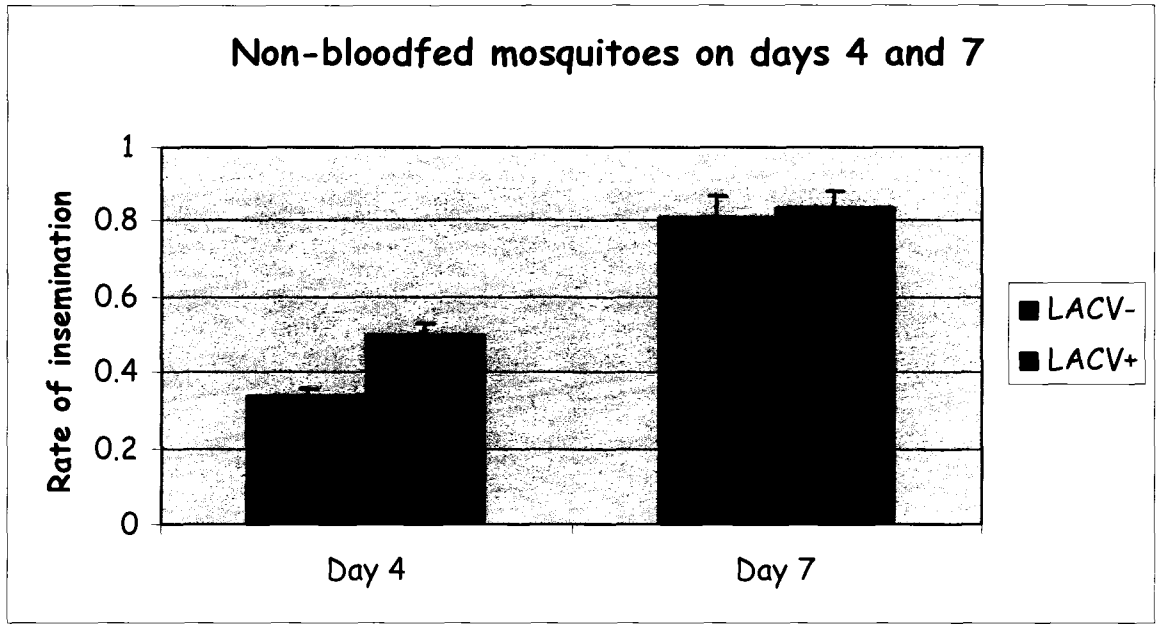



**Figure 5.2: LACV-TOT+ laboratory colonized females were inseminated more rapidly compared to uninfected mosquitoes without a bloodmeal**
The experiments were performed in triplicate and the results were averaged. There was a difference in insemination rates by day 4 (LACV+: n = 200, LACV-: n = 280) and by day 7, the insemination rates were similar for both LACV+ (n = 191) and LACV- (n = 268).

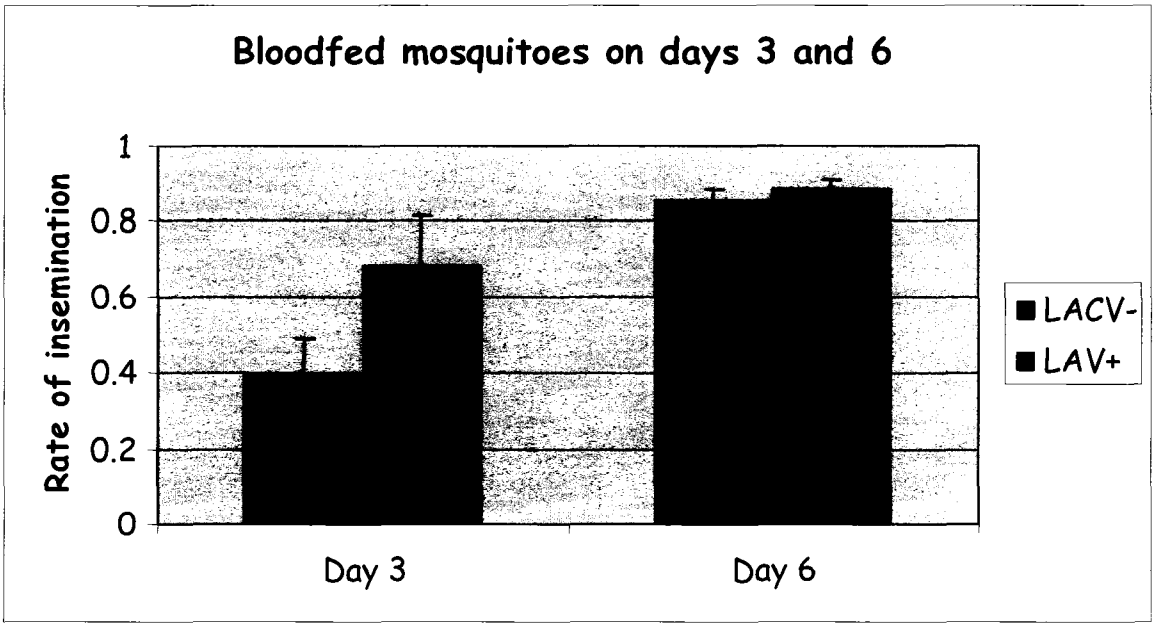



**Figure 5.3: LACV-TOT+ laboratory colonized females were inseminated more rapidly than the uninfected mosquitoes following a bloodmeal on day 3.**
The experiments were performed in triplicate and the results were averaged. The females were removed on days 3 and 6. By day 6, the insemination rate was the same for both LACV+ (n = 199) and LACV- (n = 258) mosquitoes.

33.3% (n = 24) compared to 15.7% (n = 115) for the LACV- mosquitoes ($\chi 2 = 4.08$, p = 0.043) (Figure 5.4).

**4. Comparison of insemination rates of LACV- and LACV+ field-collected and bloodfed mosquitoes**

The field collected mosquitoes were given a bloodmeal and then allowed to mate for seven days. The results in this experiment were similar to those seen with the field collected mosquitoes that did not ingest a bloodmeal. However, the differences did not differ statistically perhaps due to small sample size. The LACV+ mosquito insemination rate was 18.2% (n = 33) compared to 15.5% (n = 58) for the LACV- mosquitoes ($\chi 2 = 0.108$, p = 0.742) (Figure 5.4).

**5. Association of genome equivalents and insemination**

A portion of the infected female mosquitoes from the previous experiments were assayed by qRT-PCR to measure LACV genome equivalents and establish whether there was a relationship between LACV genome equivalents and insemination. Total RNA was extracted from the mosquitoes when the largest difference of insemination rate was observed (day 4: L+T+, non-bloodfed group, day 3: L+T+ bloodfed group, day 7: field collected bloodfed and non-bloodfed). Genome equivalents did not differ between inseminated and non-inseminated L+T+ colony mosquitoes in either the bloodfed or non-bloodfed groups. Field collected mosquitoes also showed no significant difference in LACV genome equivalents between inseminated and non-inseminated mosquitoes that were bloodfed or non-bloodfed. This suggests that the more rapid insemination is not a virus dose dependent mechanism (Table 5.2 and Figure 5.5).

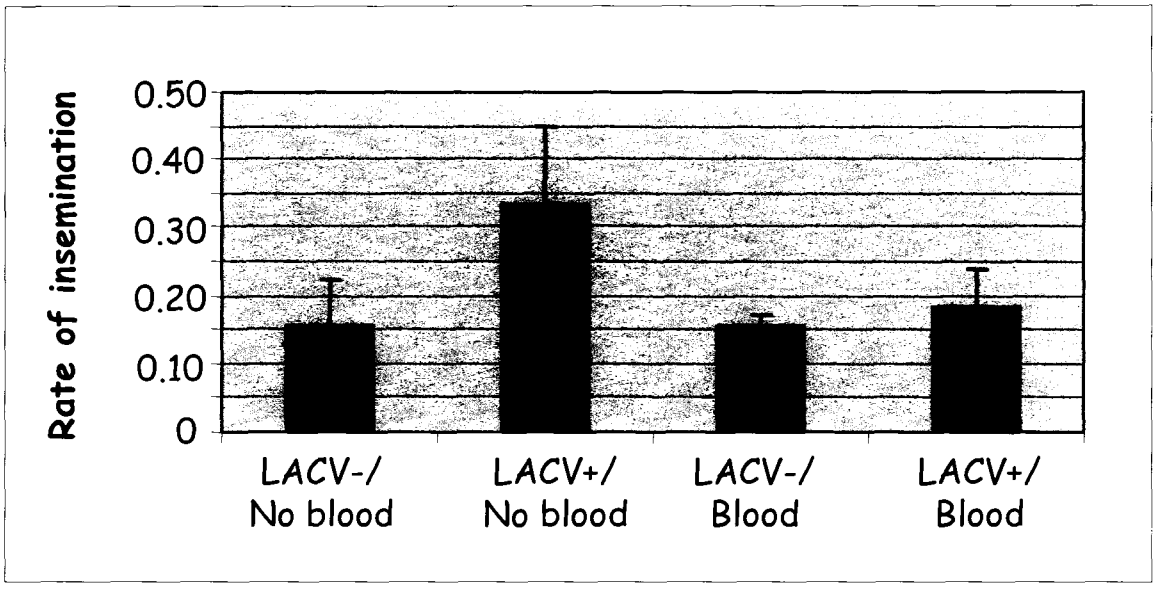



**Figure 5.4: Insemination rates were greater for LACV TOT+ field collected mosquitoes than for noninfected mosquitoes without and with a bloodmeal.**
The experiments were performed in triplicate and the results were averaged. The females were removed on day 7. The rate of insemination was higher for the LACV+ field collected mosquitoes without (LACV-: n = 115, LACV+: n = 24) and with a bloodmeal (LACV-: 58, LACV+: 33).

**Table 5.2: There is no association between genome equivalents and insemination**

| | | Average of the log of genome equivalents | |
|---|---|---|---|
| **Mosquito strain** | **Bloodfed** | **Inseminated** | **Not inseminated** |
| L+T+ strain | x | 7.22 (n = 29) | 7.03 (n = 34) |
| L+T+ strain | | 6.85 (n = 54) | 7.00 (n = 25) |
| Field-collected | x | 2.29 (n = 4) | 2.36 (n = 14) |
| Field-collected | | 2.41 (n = 6) | 2.33 (n = 18) |

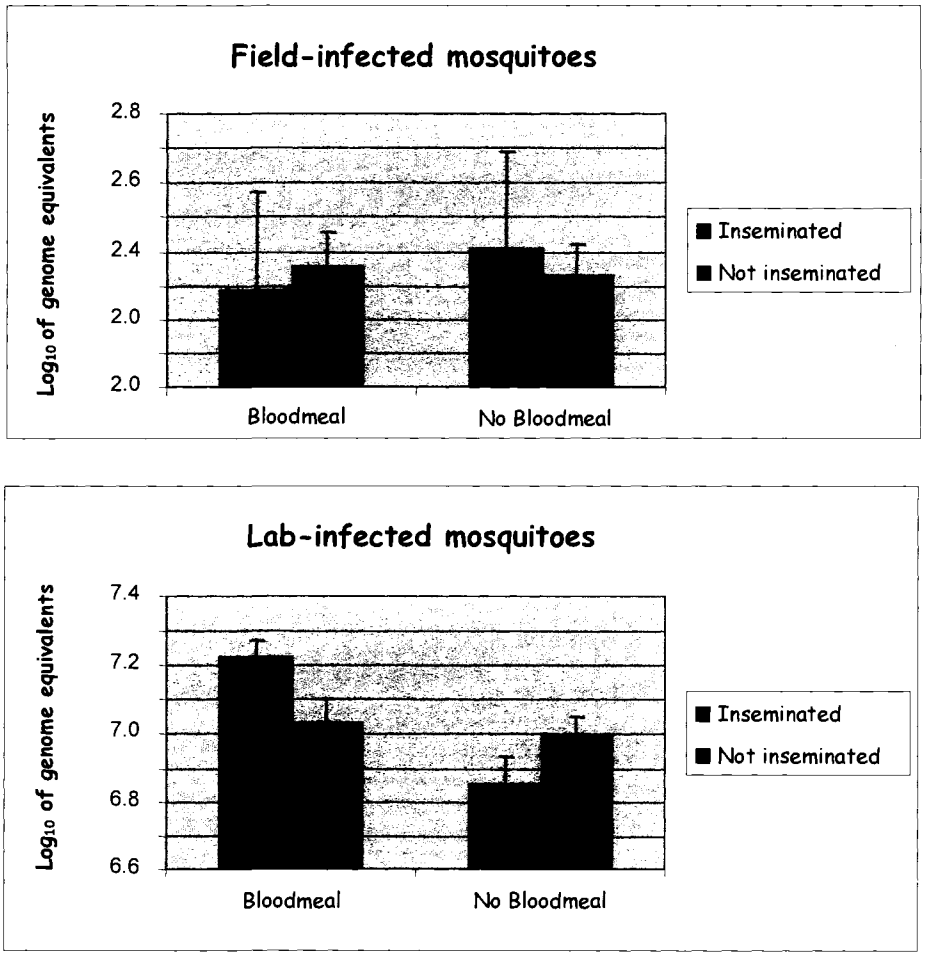



**Figure 5.5: LACV titer is not correlated with insemination efficiency** qRT-PCR was used to determine if there was a significant correlation between genome equivalents and insemination rate for both field-infected mosquitoes and the L+T+ colony mosquitoes.

**6. Cuticular Hydrocarbon Analysis**

CHC profiles were analyzed for both infected and uninfected mosquitoes. Each experiment was performed in triplicate. Nine CHCs were compared in LACV+ and uninfected mosquitoes. The main analytes were long chained saturated hydrocarbons with between 16 and 30 carbons (Figure 5.6 and Table 5.3). The primary CHCs extracted were hexadecane ($C_{16}H_{34}$), squalane ($C_{30}H_{62}$), heptadecane ($C_{17}H_{36}$), and 1-octadecanol ($C_{18}H_{38}O$). The CHC profiles of the infected and uninfected mosquitoes were similar (Figure 5.6); no new peaks or peaks with altered intensities were observed.

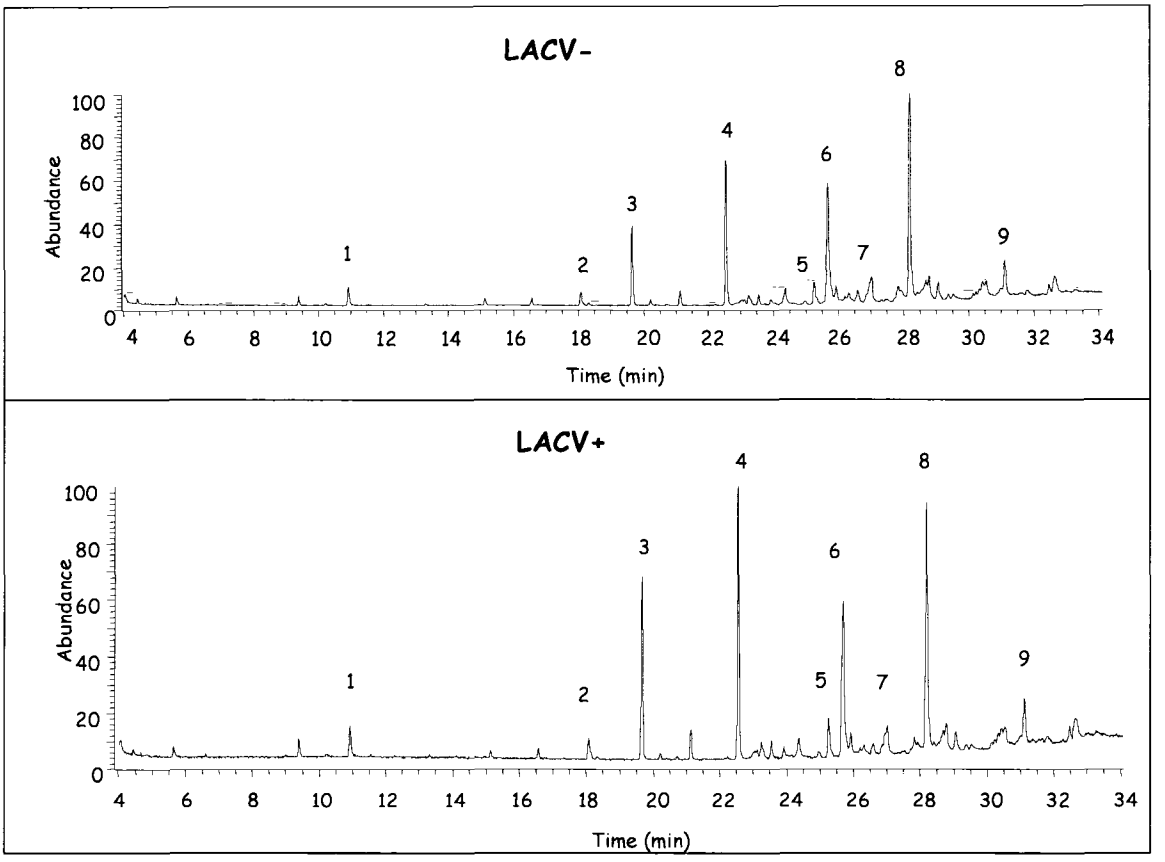



**Figure 5.6: Chromatogram of *Aedes triseriatus* mosquitoes one day post-bloodmeal** The major compounds extracted were hexadecane ($C_{16}H_{34}$) detected at 19.6 and 22.56 minutes, squalane ($C_{30}H_{62}$) detected at 25.26 minutes, heptadecane ($C_{17}H_{36}$) detected at 25.68 minutes, and 1-octadecanol ($C_{18}H_{38}O$) detected at 28.19 minutes. There was no difference in CHC profile between LACV+ and LACV-.

**Table 5.3: CHCs of adult female *Aedes triseriatus* mosquitoes**

| | | |
|---|---|---|
| 1 | 11.01 | $C_{22}H_{42}O_4$ |
| 2 | 18.07 | $C_{16}H_{34}$ |
| 3 | 19.66 | $C_{17}H_{36}O$ |
| 4 | 22.56 | $C_{16}H_{34}$ |
| 5 | 25.26 | $C_{30}H_{62}$ |
| 6 | 25.68 | $C_{17}H_{36}$ |
| 7 | 27.01 | $C_{19}H_{36}O_2$ |
| 8 | 28.19 | $C_{18}H_{38}O$ |
| 9 | 31.12 | $C_{19}H_{36}O_2$ |

**7. Female accessory sex gland**

Fourteen days post-infection, the female accessory sex gland was infected with LACV (Figure 5.7).


**Figure 5.7: LACV infection of the female accessory sex gland (100x)**

## D. DISCUSSION

These experiments clearly demonstrated that LACV infection promotes more rapid insemination of laboratory colonized transovarially-infected *Ae. triseriatus* mosquitoes (Figures 5.2-5.3). The results expanded upon the previous results demonstrating that oral LACV infection increases insemination rates in *Ae. triseriatus* mosquitoes (Gabitzsch et al., 2006). Gabitzsch *et al.* observed increased insemination rates for LACV+ mosquitoes, whereas this study observed more rapid insemination rates for the laboratory colonized LACV+ mosquitoes. The insemination rates for both infected and uninfected mosquitoes were equal by day seven (Figures 5.2-5.3). The field-infected mosquitoes had an increase in insemination rates compared to uninfected mosquitoes by day seven (Figure 5.4). These experiments demonstrated a significant association with insemination rates and LACV infection.

The mating advantage of LACV+ females observed in field-infected mosquitoes (Figure 5.4) could be very important in nature. Increasing insemination rates in infected females would increase the opportunity for TOT as well as venereal transmission of the virus, promoting virus amplification and maintenance in nature by multiple mechanisms. (Beaty et al., 2000, Gabitzsch et al., 2006). This could also compensate for the deleterious effects of LACV infection on *Ae. triseriatus* overwintering survival in natural conditions (McGaw et al., 1998).

LACV+ females could have a selective advantage in the field that would result in amplification of the virus, resulting in more infected progeny. If these offspring are transovarially-infected then the prevalence of LACV will increase in the environment. This could be very important for several reasons. In nature, more than 50% of mosquitoes

seeking a bloodmeal have not been inseminated (Porter & Defoliart, 1985, Scholl et al., 1979). Laboratory studies have shown that bloodfed females are more receptive to insemination and at an earlier time point than unfed females (Mather & Defoliart, 1984). Females given a bloodmeal had mean insemination rates of 68.2%, compared to only 49.5% of non-bloodfed mosquitoes by day four post-mixing with males (Figures 5.2 and 5.3). Perhaps in the field the LACV+ bloodfed females may be more rapidly inseminated. The more rapid insemination of LACV+ mosquitoes seen in laboratory cage studies would likely be amplified in importance in forested natural areas where mate seeking would be much more complex.

The insemination rates between infected and uninfected bloodfed field-collected mosquitoes did not differ statistically, although the infected were inseminated (18.2%) more often than the uninfected (15.5%) (Figure 5.4). This particular experiment was complicated by the small sample size of field-infected mosquitoes. Only 3-5% of collected mosquitoes are infected with LACV (II. C1, IV. C1), which made it difficult to obtain a large sample size. The small sample size was also due to the small number of females that were mated. The small numbers of LACV+ females in the cage for each trial (5-8) and infrequent mating made it difficult to observe a significant effect. A greater sample size could have provided statistically significant results.

LACV infection probably promotes mating efficiency through either a physical or behavioral mechanism. In this study, three possible mechanisms that could explain the more rapid insemination were investigated. They involved individual mosquito virus titers, CHC profiles, and accessory sex gland infection. Differences in individual mosquito virus titers could have a significant effect on transmission and behavior and

therefore was investigated through qRT-PCR (Watts et al., 1972). A dose dependency of infection has been observed with Sigma virus and *Drosophila melanogaster* (Seecof, 1966). Flies transmitting Sigma virus to relatively few offspring can contain higher virus titers than flies transmitting the infection to nearly all their progeny. High virus titers are associated with deaths among a particular strain of *Drosophila melanogaster* (Seecof, 1964). However qRT-PCR analyses determined that the more rapid insemination was not a dose-dependent mechanism (Table 5.2 and Figure 5.5).

The female accessory gland could possibly play a role in the observed mating advantage. The accessory gland is located behind the spermathecal openings. In some insects, the accessory gland secretes a cement for attaching the eggs to a support or a material to form an egg covering (Snodgrass, 1959). It has also been demonstrated that in *Musca domestica* the sperm are either "activated" or the permeability of the egg membrane is altered by the secretion of the accessory sex glands before fertilization occurs (Leopold & Degrugillier, 1973). The role of the female accessory gland in the mating rituals of mosquitoes is not understood, but it is intriguing that the accessory glands of LACV+ *Ae. triseriatus* mosquitoes are infected (Figure 5.7). The accessory glands could play a role in pheromone production resulting in a more attractive female. The accessory glands could change or increase the production of the sex pheromones and therefore influencing the mating efficiency.

Stimulatory pheromones have an important role during the mating ritual of the order Diptera (Ferveur et al., 1996, Ferveur et al., 1997, Greenspan & Ferveur, 2000, Jallon, 1984). They are volatile and would be detected at a distance of less than a few centimeters before the first physical contact (Greenspan & Ferveur, 2000). Many

dipterans use sex pheromones with CHC components for recognition and these substances often are altered with physiological state and age (Pomonis, 1989, Trabalon et al., 1988). Recent evidence has suggested that mating alters the CHC profile of female *Anopheles gambiae* and *Ae. aegypti* (Polerstock et al., 2002). In *A. gambiae*, the proportions of two CHC components, *n*-heneicosane and *n*-tricosene, were significantly reduced as the female aged and mated. Female *Ae. aegypti* showed significant changes in the proportions of *n*-heptadecane, *n*-pentacosane, and *n*-hexacosane in their cuticles after mating again suggesting the importance of CHCs in chemical communication during mosquito courtship (Polerstock et al., 2002).

The CHCs of LACV+ *Ae. triseriatus* females were investigated to determine if LACV alters the CHC profiles, potentially resulting in infected females becoming more attractive than uninfected ones and promoting more rapid insemination. Ovarian-produced ecdysteroids control the modification of CHCs making them more attractive to males. It is possible that LACV infection of the ovaries could speed up this process.

Although it is an attractive hypothesis that CHCs could be involved in more rapid insemination of LACV+ females, the analysis showed no difference in the CHC profiles between LACV+ and uninfected females. This could be because of technical issues. The CHC analysis should have been completed with a known internal standard to better confirm changes in CHC profiles. Since this analysis was not complete, it is still possible that a change in CHC profile could contribute to the more rapid insemination observed with LACV+ mosquitoes. This experiment should be repeated using an internal standard to detect CHC changes in the different mosquito groups and more accurately determine if

changes in CHC profiles really are associated with more rapid insemination of LACV+ females.

Numerous additional factors could be affected by LACV infection resulting in faster insemination. The female mosquito must endure physiological changes to become prepared for mating, including alteration in wing beat and CHC production. Female mosquitoes control the refractory period allowing neither the physical coupling nor the cues that are necessary for male ejaculation (Klowden & Zwiebel, 2005). A post-eclosion production of juvenile hormone (JH) appears to cause the development of mating competence. It is possible that the release of JH in a LACV+ mosquito could occur earlier than in an uninfected mosquito allowing the LACV+ females to be receptive to mating faster than uninfected females. There are a number of factors influencing mating of mosquitoes including acoustics, chemical cues and behavior. Any one of these factors or a combination of these factors could be affected by a LACV infection. It is provocative to hypothesize that LACV infection may alter the female mosquitoes to become either more attractive or receptive to mating more quickly than uninfected mosquitoes. Future studies will need to be done to explore the association between LACV infection and insemination.

The observation that insemination is occurring more rapidly in infected females compared to uninfected females could have major epidemiological implications. Transovarially infected females that emerge and immediately mate will then produce infected progeny to continue the cycle. Over an entire summer, if LACV+ females are mating more rapidly than uninfected, even if the overall rate of insemination is the same between the two groups, the number and percentage of infected mosquitoes will increase.

As previously discussed in Chapter IV, mathematical models of LACV have suggested that field infection rates are too low for LACV to persist in nature. These models have not taken into account the consideration of a possible mating advantage associated with LACV infection. The mating advantage could contribute to LACV persistence in nature and promote virus amplification and maintenance.

# VI. Summary and Discussion

## I. SUMMARY AND DISCUSSION

Arthropod-borne viruses (arbovirus) are emerging and resurging worldwide. There are many factors contributing to the resurgence and emergence of arboviruses. These factors include evolution of arboviruses through genetic drift and segment reassortment and increased air travel, urbanization, and population. La Crosse virus (LACV) is a prototypical emerged arbovirus. In this dissertation the evolutionary, epidemic and maintenance potential of LACV in nature were investigated.

There has been ample evidence of evolution of viruses in the *Bunyaviridae* family in nature through both genetic drift (mutation) and genetic shift (segment reassortment). LACV has demonstrated significant evolutionary potential through segment reassortment in both field and laboratory studies. *Aedes triseriatus* mosquitoes are capable of dual infection resulting in segment reassortment, which has been demonstrated in many laboratory studies. The evolutionary potential of LACV was explored and significant genetic variation was observed in LACV RNA amplified from field collected mosquitoes suggesting frequent segment reassortment of LACV in nature. A maximum parsimony phylogenetic analysis and a linkage disequilibrium analysis revealed 25-38.6% of the samples were reassortants. There was also evidence of quasispecies in some of the mosquitoes analyzed, which may confound the results but demonstrates the evolutionary potential of LACV. Further studies need to be performed to determine if infectious virus has genome segments demonstrating quasispecies-like sequence variability. Viruses will need to be isolated from mosquitoes and plaque-purified clones will need to be sequenced to determine if this variability is indeed occurring in nature.

There are epidemiological implications of segment reassortment in LACV+ *Ae. triseriatus* mosquitoes in the field. If a new virus is created, important determinants of its pathogenesis and transmission cycle could be altered. New viral phenotypes could be capable of infecting new vector species or vertebrate hosts, and new viruses could be introduced into new arbovirus cycles. This could potentially have significant epidemiological consequences. If a new viral genotype established a new transmission cycle with a different mosquito that is more aggressively anthropophilic, increased human infection could occur, possibly becoming clinically significant in both adults and children. In addition, a new viral phenotype could become more virulent and exhibit different tropisms perhaps causing humans to develop a high enough viremia to infect biting mosquitoes. Determination of the evolutionary potential of LACV through segment reassortment may permit prediction of the epidemiologic consequences of those events.

The geographical, environmental and temporal factors that condition the genetic structure of LACV were also investigated. Analysis of LACV amplified from field-infected mosquitoes collected in 2000, 2004, and 2005, revealed that there were no barriers to viral gene flow in the study area, no isolation by distance and therefore, all the LACV S and M segments can be expected to circulate effectively throughout the study area. This has epidemiological implications as well. Previous studies have suggested that there may be a narrow range of genotypes correlated with fatal outcomes. The presence and potential trafficking of such genotypes could pose major risks to humans. In this study, all the LACV strains, regardless of virulence, could traffic and be transmitted throughout the entire 15,360 $km^2$ study range.

Although there were no barriers to viral gene flow or isolation by distance, a very significant temporal association was revealed. Defined haplotypes for the S (5 haplotypes) and M (6 haplotypes) segments were identified and the haplotypes clustered by collection year. The variation observed in the M segment is highly suggestive of diversifying selection due to immune factors, specifically neutralizing antibodies from the vertebrate host possibly selecting for new haplotypes each year. The variation in the S segment could involve linkage between gene products. In segmented viruses, all segments contain sequences important for packaging and the temporal determinants may be explained by linkage equilibrium due to the interaction between the nucleoprotein and the glycoprotein during packaging. The variability observed in this analysis may, once again, be a result of quasispecies. In order to determine if the temporal association with genetic variation is indeed occurring in the field, virus isolation, plaque purification and sequencing needs to be performed. However, passage *in vitro* may also reduce the observed variability. If the temporal association is occurring in the field, control by vaccination may be difficult. For example, with influenza virus, a new vaccine is developed each year to counter the temporal variability.

The maintenance of LACV in nature is not well understood. Mathematical models have revealed that field infection rates are well below the rates needed to maintain the virus in nature. However, the mathematical models have not considered the possibility of stably-infected *Ae. triseriatus* mosquitoes. High-titered, super-infected field collected mosquitoes were discovered in collections 2006 and 2007 at a low rate. In total, 0.06% of (16/19,471) of all mosquitoes assayed for LACV were super-infected. The identification of super-infected mosquitoes suggests that LACV could be maintained in nature through

a stabilized infection in a small number of females. If the super-infected mosquitoes transmit LACV to all of most of their progeny, it is conceivable that the virus could be maintained in nature at a fairly constant level by a relatively small number of stably infected females. Further studies need to determine is the super infection is due to mosquito genetics, viral genetics or possible both. Analyses would include determination of TOT potential of super-infected mosquitoes, comparison of gene expression of super-infected to normally-infected mosquitoes, microarray analyses, genetic crossing experiments, and mosquito pathogenesis. The potential effect of super-infected mosquitoes on LACV maintenance in nature needs to be elucidated.

Multiple phylogenetic analyses with five virus isolates isolated from the super-infected mosquitoes and previously published LACV sequences revealed that three of the isolates from three different collection sites were genetically similar, with few polymorphisms. Two isolates from super-infected mosquitoes were from the same collection site, collected in 2006 and 2007, were genetically similar to each other but differed from other viruses isolated from super-infected mosquitoes suggesting purifying selection within a collection site. In the M segment phylogenetic analysis, one virus isolated from a super-infected mosquito clustered with a virus isolate from a known human fatality. This clustering could suggest a possible association with super-infected mosquitoes and human fatality. Perhaps the virus causing super infections in mosquitoes may cause severe clinical outcomes in humans. This could be due to dose inoculum of the saliva, since the average titer of the super-infected mosquitoes is extremely high or to linkage of virus replication and persistence in mosquitoes and efficiency with human

replication. Future studies need to be performed with the LACV isolates obtained from super-infected mosquitoes in order to determine genetic determinants of super infection.

LACV maintenance in nature may also be assisted by a mating advantage confirmed by LACV infection in female *Ae. triseriatus* mosquitoes. This is an additional factor that mathematical models have not considered. LACV transovarially-infected female mosquitoes become inseminated more rapidly than uninfected mosquitoes. The mating advantage associated with LACV infection observed in field-infected female mosquitoes could be very important in nature. Increasing insemination rates in infected females would increase the chance for transovarial transmission as well as venereal transmission of the virus. This would promote virus amplification and maintenance in nature by multiple mechanisms. The mechanism of the mating advantage has yet to be determined. There are a number of factors influencing mating of mosquitoes and any one of the factors or a combination of these factors could be affected by a LACV infection. It is provocative to think that LACV infection possibly alters the female mosquitoes to become either more attractive or receptive to mating more quickly than uninfected mosquitoes and thus promotes virus amplification and maintenance in nature.

The evolutionary potential of LACV has major epidemiological implications and consistently needs to be examined. These studies have shown that LACV is constantly evolving either through genetic drift or more importantly, segment reassortment. The constant evolution is especially important since on a small spatial scale there are no barriers to viral flow and no isolation by distance.

Persistence of LACV in nature has not been well understood, however, stabilized infection and mating advantage for infected mosquitoes may certainly contribute to

LACV maintenance in nature. These investigations of the evolutionary and maintenance potential of LACV can provide insight into the determinants of arbovirus emergence and epidemic potential.

# APPENDIX 1

## Appendix 1: S segment sequence alignment of LACV prototype and virus isolated from super-infected mosquitoes

```
MN/Human/1960     AGTAGTGTAC TCCACTTGAA TACTTTGAAA ATAAATTGTT GTTGACTGTT TTTTACCTAA GGGGAAATTA TCAAGAGTGT GATGTCGGAT TTGGTGTTTT   100
BEN2/Mosq/2007    T......... C......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     T......... C......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  T......... C......... .......... .......... .......... .......... .......... .T........ .......... ..........
NAT/Mosq/2006     T......... C......... .......... .......... .......... .......... .A........ .......... .......... ..........
NAT/Mosq/2007     T......... C......... .......... .......... .......... .......... .A........ .......... .......... ..........

MN/Human/1960     ATGATGTCGC ATCAACAGGT GCAAATGGAT TTGATCCTGA TGCAGGGTAT ATGGACTTCT GTGTTAAAAA TGCAGAATCA CTCAACCTTG CTGCAGTTAG   200
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

MN/Human/1960     GATCTTCTTC CTCAATGCCG CAAAGGCCAA GGCTGCTCTC TCGCGTAAGC CAGAGAGGAA GGCTAACCCT AAATTTGGAG AGTGGCAGGT GGAGGTTATC   300
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

MN/Human/1960     AATAATCATT TTCCTGGAAA CAGGAACAAC CCAATTGGTA ACAACGATCT TACCATCCAC AGATTATCTG GGTATTTAGC CAGATGGGTC CTTGATCAGT   400
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........

MN/Human/1960     ATAACGAGAA TGATGATGAG TCTCAGCACG AGTTGATCAG AACAACTATT ATCAACCCAA TTGCTGAGTC TAATGGTGTG GGATGGGACA GTGGGCCAGA   500
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... C......... .......... ..........
NAT/Mosq/2006     .......... ......C... .......... ....A..... .......... .......... .......... .......... .......... ....G.....
NAT/Mosq/2007     .......... ......C... .......... ....A..... .......... .......... .......... .......... .......... ....G.....
```

```
MN/Human/1960       GATCTATCTA TCATTCTTTC CAGGAACAGA AATGTTTTTG GAAACTTTCA AATTCTACCC GCTGACCATT GGAATTCACA GAGTCAAGCA AGGCATGATG   600
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .....T..C. .......... .......... .......... .......... ......T... .......... .......... ..........
NAT/Mosq/2007       .......... .....T..C. .......... .......... .......... .......... ......T... .......... .......... ..........

MN/Human/1960       GACCCTCAAT ACCTGAAGAA GGCCTTAAGG CAACGCTATG GCACTCTCAC AGCAGATAAG TGGATGTCAC AGAAGGTTGC AGCAATTGCT AAGAGCCTGA   700
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       ........G. .......... .......... .......... ....C..... .......... .......... .......... .......... ..........
NAT/Mosq/2007       ........G. .......... .......... .......... ....C..... .......... .......... .......... .......... ..........

MN/Human/1960       AGGATGTAGA GCAGCTTAAA TGGGGAAAAG GAGGCCTGAG CGATACTGCT AAAACATTCC TGCAGAAATT TGGCATCAGG CTTCCATAAA TATGGCATGA   800
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... ....A..... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......A.. .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ...A......
NAT/Mosq/2006       .......... .......... .......... .......... .......... .......... ....A..... .......... ..C....G.. ..G.......
NAT/Mosq/2007       .......... .......... .......... .......... .......... .......... ....A..... .......... ..C....G.. ..G.......

MN/Human/1960       GGCATTCAAA TTAGGTTCTA AATTCTAAAT TTATATATGT CAATTTGATT AATTGGTTAT CCAAAAGGGT TTTCTTAAGG GAACCCACAA AAATAGCAGC   900
BEN2/Mosq/2007      ........G. .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... ....A..... .......... .......... T......... .......... .......... .......... .......... ..........
NAT/Mosq/2007       .......... ....A..... .......... .......... T......... .......... .......... .......... .......... ..........

MN/Human/1960       TAAATGGGTG GGTGGTAGGG GACAGCAAAA AACTATAAAT CAGGTCATAA ATAAAATAAA ATGTATTCAG TGGAGCACAC TACT     984
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... ...G...... ...A
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... ...G...... ...A
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... ...G...... ...A
NAT/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... ...G...... ...A
NAT/Mosq/2007       .......... .......... .......... .......... .......... .......... .......... ...G...... ...A
```

# APPENDIX 2

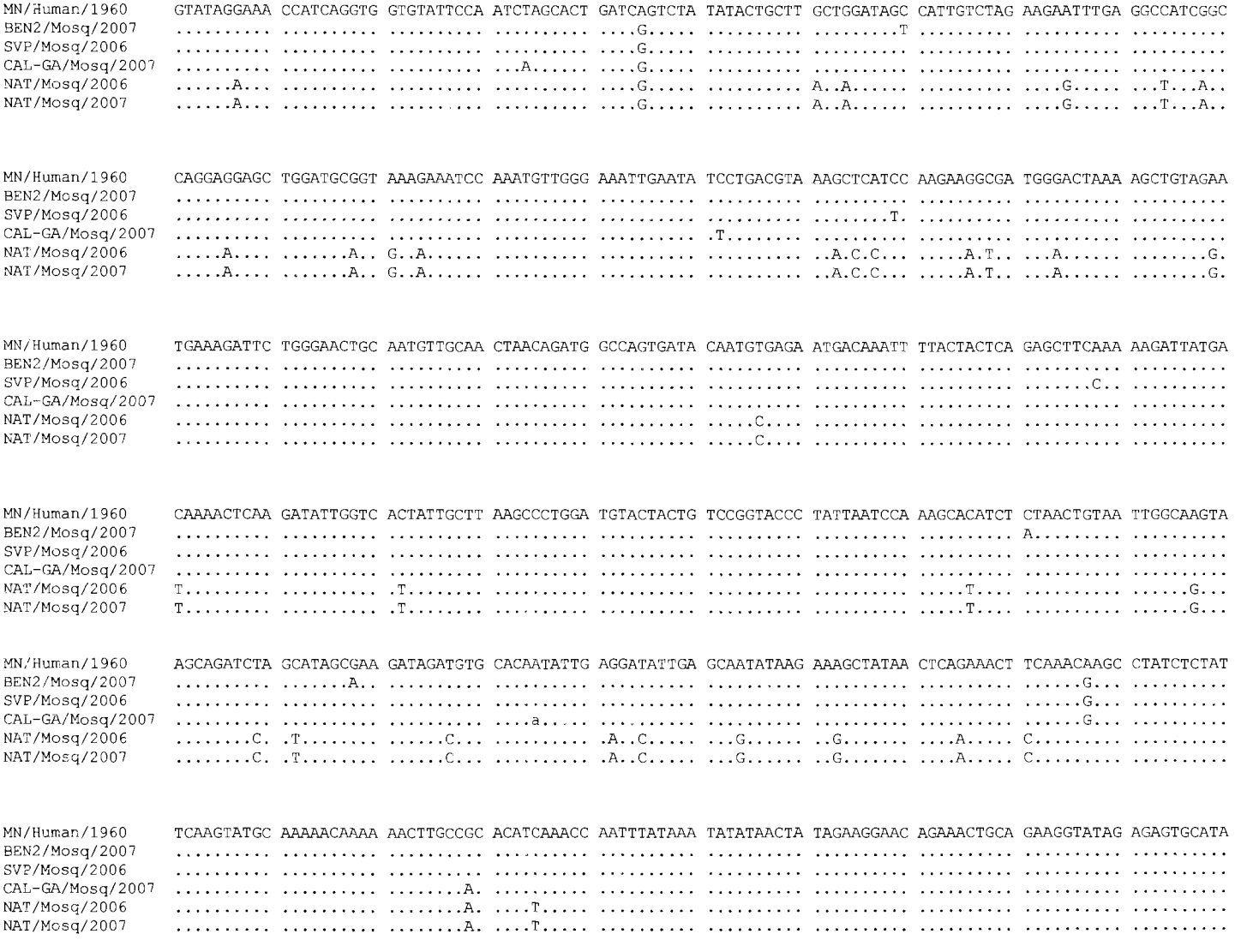
# Appendix 2: M segment sequence alignment of LACV prototype and viruses isolated from super-infected mosquitoes

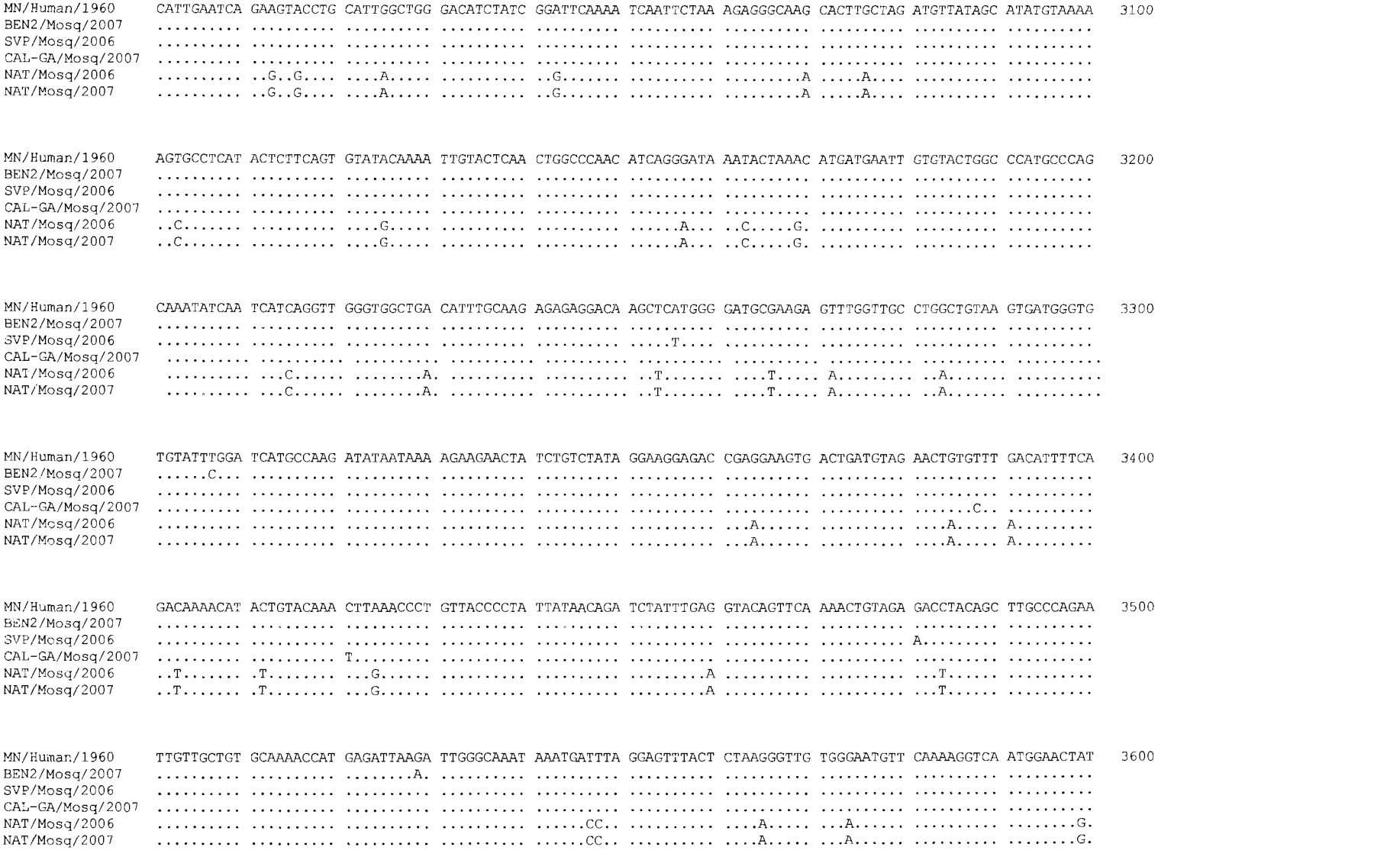
```
MN/Human/1960      AGTAGTGTAC TACCAAGTAT AGATAACGTT TGAATATTAA AGTTTTGAAT CAAAGCCAAA GATGATTTGT ATATTGGTGC TAATTACAGT TGCAGCTGCA    100
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .G........ ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007      .......... .......... .......... .......... .A........ .......... .......C.C ...C...... .G...G.... .A........

MN/Human/1960      AGCCCAGTGT ATCAAAGGTG TTTCCAAGAT GGGGCTATAG TGAAGCAAAA CCCATCCAAA GAAGCAGTTA CAGAGGTGTG CCTGAAAGAT GATGTTAGCA    200
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007      .......... .......A.. .......... .......... .......... .......... ........C. .......... .......... ..........

MN/Human/1960      TGATCAAAAC AGAGGCCAGG TATGTAAGAA ATGCAACAGG AGTTTTTTCA AATAATGTCG CAATAAGGAA ATGGCTAGTC TCTGATTGGC ATGATTGCAG    300
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... ...C...... .......... .......... .......... .......... ..........
NAT/Mosq/2006      ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007      .......... .......... .......... .......... G......... ..C.....T. .......... .......... .......... ..........

MN/Human/1960      GCCTAAGAAG ATCGTTGGGG GACACATCAA TGTAATAGAA GTTGGTGATG ACCTGTCACT CCATACTGAA TCATATGTTT GCAGCGCAGA TTGTACCATA    400
BEN2/Mosq/2007     .......... .......... .......... .......... ........C. .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007      .......... ...A...... .......... .......... .......... ....A..G.. G......... .......... .......... ......A...

MN/Human/1960      GGTGTAGACA AAGAGACTGC ACAGGTCAGG CTTCAGACAG ATACCACAAA TCATTTTGAA ATTGCAGGCA CTACTGTGAA GTCAGGATGG TTCAAGAGCA    500
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007      .......... .......... G........A .......... ....T..... .......... .......... .......A.. .......... .....A....

MN/Human/1960      CGACATATAT AACTCTTGAT CAAACTTGCG AACACCTTAA AGTTTCCTGC GGCCCAAAAT CTGTACAGTT CCATGCCTGC TTCAATCAGC ATATGTCTTG    600
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007      .......... .......... .....C.... .......... .........T .......... .......A.. .......... .......... ..........
```

```
MN/Human/1960     CGTCAGATTT TTACACAGGA CAATATTGCC TGGCTCTATA GCCAATTCCA TATGTCAGAA TATCGAAATC ATAATTTTAG TTACACTTAC TCTATTAATC     700
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... ....T..... .......... ........T. .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .......... .........A. .......... ......C..T .......... .......A.. ...T...... .......... .......... ......G...


MN/Human/1960     TTTATATTGT TAAGCATTTT AAGTAAGACT TATATATGTT ATTTATTAAT GCCTATATTC ATCCCCATAG CATATATATA CGGTACAATT TACAATAAGT     800
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .....T.... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .....T.... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .....T.... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .......... ....T..... .......... .......... .......... .........T .......... .....G.... .....T.... ..........


MN/Human/1960     CGTGCAAAAA ATGCAAATTA TGTGGCTTAG TGTATCACCC ATTCACAGAG TGTGGCACAC ATTGTGTCTG TGGTGCCCGC TATGATACTT CAGATAGAAT     900
BEN2/Mosq/2007    .......... .......... .......... .......T.. .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......T.. .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .T........ .......... .......... .......T.. .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .......... .......C.. ..C...C... .......T.. .......... .......... .......... .......... .......... ..........


MN/Human/1960     GAAACTGCAT AGAGCTTCTG GATTGTGCCC TGGTTATAAA AGCCTAAGAG CTGCCAGAGT CATGTGCAAG TCGAAAGGGC CTGCATCAAT ATTGTCTATA    1000
BEN2/Mosq/2007    .........C .......... ....A..... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... ....A..... .......... .......... ....T..... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .........C .......... .......... .......... .......... .......... .......... .......... .......... T..A......


MN/Human/1960     ATTACTGCGG TACTGGTCTT AACCTTTGTG ACACCAATCA ACTCCATGGT TTTAGGAGAG AGTAAAGAAA CCTTTGAACT TGAAGATCTT CCAGACGACA    1100
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     ........A. .......... .........A .......... ....T..... ......G..A .......... .G........ ......G... .....T....


MN/Human/1960     TGTTGGAAAT GGCATCGAGA ATAAATTCTT ATTATCTTAC CTGTATCTTG AATTATGCTG TAAGCTGGGG TCTTGTTATC ATTGGATTGT TGATCGGGCT    1200
BEN2/Mosq/2007    .......... .......... .......... .......C.. .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......C.. .......... .......... .......... .......... .......... .....G....
CAL-GA/Mosq/2007  .......... .......... .......... .......C.. .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .......... .........G .......... .....T.... ...C...... .......... .......... .......... .....G.... ...C...A..
```

```
MN/Human/1960     GCTTTTTAAG AAATACCAGC ACAGATTCTT AAATGTTTAC GCAATGTACT GTGAAGAATG TGACATGTAT CATGACAAGT CTGGGTTGAA AGGACATGGT    1300
BEN2/Mosq/2007    ...A...... .......... .......... .......... .......... .......... .......... .......... .......... .A.......
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... .A.......
CAL-GA/Mosq/2007  .......A.. .......... ......T..T .....T.... .......... .......... .......... .......... .......... .A.G......
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     ...A..C... ........A. .......T.. .........T .......... ....G..... .......... .......... .......A.. .A.......

MN/Human/1960     GATTTCACCA ACAAATGCAG ACAGTGCACA TGTGGTCAAT ATGAAGATGC TGCAGGTTTG ATGGCTCACA GGAAAACCTA TAACTGCTTA GTGCAGTACA    1400
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     ..C..T.... .......... G.....T... .....C.... .......... CA....C... .....C.... .......... .......... ..........

MN/Human/1960     AAGCAAAGTG GATGATGAAC TTCCTGATAA TTTACATATT CTTAATTTTG ATCAAAGATT CTGCTATAGT TGTACAAGCT GCTGGAACTG ACTTCACCAC    1500
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ...T......
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .......... .........T .....A.... .......... .........A .......... .......... .......... ........C. .T.......

MN/Human/1960     CTGCCTAGAG ACTGAGAGTA TAAATTGGAA CTGCACTGGG CCATTTTTGA ACCTCGGGAA TTGCCAAAAG CAACAAAAGA AAGAACCTTA CACCAACATT    1600
BEN2/Mosq/2007    .......... .......... .......... .......... ..G....... .......T.. .......... .......... .......... ..........
SVP/Mosq/2006     .......... ........C. .......... .......... .......... .......... .........A .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     T......... .......... .......... .......... ..G....... .......... ...T...... ..G....... .G........ ..........

MN/Human/1960     GCAACTCAGT TAAAGGGACT AAAGGCAATT TCCGTACTAG ATGTCCCTAT AATAACAGGG ATACCAGATG ATATTGCGGG TGCTTTAAGA TATATAGAAG    1700
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... ....C..... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .....C..A. .......G.. .......... ........G. ..A....... .......... ..T....... .......A.. .........G ..........
MN/Human/1960     AGAAGGAAGA TTTCCATGTC CAGCTAACTA TAGAATATGC GATGTTAAGT AAATACTGTG ACTATTATAC CCAATTCTCA GATAACTCAG GATACAGTCA    1800
BEN2/Mosq/2007    ....A..... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... ..G....... .......... ...A...... .........G ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .A.G...... .......... .......... .......C.. A...C.G... ..G....... .......C.. .......... .......... ..........
```

```
MN/Human/1960     GACAACATGG AGAGTGTACT TAAGGTCTCA TGATTTTGAA GCCTGTATAC TATATCCAAA TCAGCACTTT TGCAGATGTG TAAAAAATGG TGAGAAGTGC   1900
BEN2/Mosq/2007    .......... .......... .......... ...C...... .......... .......A.. .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......A.. .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......A.. .....G.... .......... C.....G... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .G........ ..G....... .......... .........G ..A....... .......... ......T... .......... .G........ ...A......

MN/Human/1960     AGCAGCTCCA ATTGGGACTT TGCCAATGAA ATGAAAGATT ATTACTCTGG GAAACAAACA AAGTTTGACA AGGATTTAAA TCTAGCCCTA ACAGCTTTGC   2000
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... ....C..... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......G.. .......... ....C..... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... ....C..... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007     .......... .......T.. .......... .....G.... T......... A......... ........T. .A..C..... ...G...... ..........

MN/Human/1960     ATCATGCCTT CAGGGGGACC TCATCTGCAT ATATAGCAAC AATGCTCTCA AAAAAGTCCA ATGATGACTT GATTGCATAC ACAAATAAGA TAAAAACAAA   2100
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... ........?. ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... ........?. ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ?????????? ?????????? .......... .......... .......... .......... .......... .......... ........A. ..........
NAT/Mosq/2007     .C........ .......... .......... .......... .......... .......... .......... .......... .......... ..........

MN/Human/1960     ATTCCCAGGT GATGCATTGT TGAAGGCTAT AATAGATTAT ATAGCATATA TGAAAAGTTT GCCAGGTATG GCAAATTTCA AATATGATGA ATTCTGGGAT   2200
BEN2/Mosq/2007    .......... A......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... A......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... A......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ...T...... A......... .......... .......... .......... .......... .......... .......... .......... G.........
NAT/Mosq/2007     ...T...... A......... .......... .......... .......... .......... .......... .......... .......... G.........

MN/Human/1960     GAATTATTGT ACAAACCCAA CCCAGCAAAG GCCTCAAACC TTGCTAGAGG AAAGGAGTCA TCTTACAACT TCAAACTAGC AATTTCATCA AAGTCTATAA   2300
BEN2/Mosq/2007    .......... .......... T......... .......... ...T...... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .........A .......... .......... .........T .......... .T........ .......... ..A.......
NAT/Mosq/2007     .......... .......... .........A .......... .......... .........T .......... .T........ .......... ..A.......

MN/Human/1960     AAACCTGCAA GAATGTTAAG GATGTTGCCT GCTTATCGCC AAGGTCAGGT GCTATATATG CTTCAATAAT TGCGTGTGGT GAACCCAATG GGCCAAGTGT   2400
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... ...A...... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .....C.... .......A.. .......... .......... ........a. .........C .......... ..........
NAT/Mosq/2007     .......... .......... .....C.... .......A.. .......... .......... .......... .........C .......... ..........
```


MN/Human/1960 GTATAGGAAA CCATCAGGTG GTGTATTCCA ATCTAGCACT GATCAGTCTA TATACTGCTT GCTGGATAGC CATTGTCTAG AAGAATTTGA GGCCATCGGC 2500
BEN2/Mosq/2007 ..... ..... ..... ....G..... ..... .....T..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ....G..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ...A..... ....G..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ......A... ..... ..... ..... ....G..... ..... A..A..... ..... ....G..... ...T...A..
NAT/Mosq/2007 ......A... ..... ..... ..... ....G..... ..... A..A..... ..... ....G..... ...T...A..
MN/Human/1960 CAGGAGGAGC TGGATGCGGT AAAGAAATCC AAATGTTGGG AAATTGAATA TCCTGACGTA AAGCTCATCC AAGAAGGCGA TGGGACTAAA AGCTGTAGAA 2600
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... .....T. ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... .T..... ..... ..... ..... .....
NAT/Mosq/2006 .....A.... .....A.. G..A..... ..... ..... ..... ..A.C.C... .....A.T.. ...A..... .....G.
NAT/Mosq/2007 .....A.... .....A.. G..A..... ..... ..... ..... ..A.C.C... .....A.T.. ...A..... .....G.
MN/Human/1960 TGAAAGATTC TGGGAACTGC AATGTTGCAA CTAACAGATG GCCAGTGATA CAATGTGAGA ATGACAAATT TTACTACTCA GAGCTTCAAA AAGATTATGA 2700
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... .....C.. .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... .....C.... ..... ..... ..... .....
NAT/Mosq/2007 ..... ..... ..... ..... ..... .....C.... ..... ..... ..... .....
MN/Human/1960 CAAAACTCAA GATATTGGTC ACTATTGCTT AAGCCCTGGA TGTACTACTG TCCGGTACCC TATTAATCCA AAGCACATCT CTAACTGTAA TTGGCAAGTA 2800
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... A..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 T..... ..... .T..... ..... ..... ..... ..... .....T.... ..... ......G...
NAT/Mosq/2007 T..... ..... .T..... ..... ..... ..... ..... .....T.... ..... ......G...
MN/Human/1960 AGCAGATCTA GCATAGCGAA GATAGATGTG CACAATATTG AGGATATTGA GCAATATAAG AAAGCTATAA CTCAGAAACT TCAAACAAGC CTATCTCTAT 2900
BEN2/Mosq/2007 ..... .....A.. ..... ..... ..... ..... ..... ..... ......G... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ......G... .....
CAL-GA/Mosq/2007 ..... ..... ..... ....a..... ..... ..... ..... ..... ......G... .....
NAT/Mosq/2006 .....C. .T..... .....C... ..... .A..C..... ...G..... ..G..... ....A..... C..... .....
NAT/Mosq/2007 .....C. .T..... .....C... ..... .A..C..... ...G..... ..G..... ....A..... C..... .....
MN/Human/1960 TCAAGTATGC AAAAACAAAA AACTTGCCGC ACATCAAACC AATTTATAAA TATATAACTA TAGAAGGAAC AGAAACTGCA GAAGGTATAG AGAGTGCATA 3000
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... .....A. ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... .....A. ....T..... ..... ..... ..... ..... .....
NAT/Mosq/2007 ..... ..... .....A. ....T..... ..... ..... ..... ..... .....


MN/Human/1960 CATTGAATCA GAAGTACCTG CATTGGCTGG GACATCTATC GGATTCAAAA TCAATTCTAA AGAGGGCAAG CACTTGCTAG ATGTTATAGC ATATGTAAAA 3100
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 AGTGCCTCAT ACTCTTCAGT GTATACAAAA TTGTACTCAA CTGGCCCAAC ATCAGGGATA AATACTAAAC ATGATGAATT GTGTACTGGC CCATGCCCAG 3200
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 CAAATATCAA TCATCAGGTT GGGTGGCTGA CATTTGCAAG AGAGAGGACA AGCTCATGGG GATGCGAAGA GTTTGGTTGC CTGGCTGTAA GTGATGGGTG 3300
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 TGTATTTGGA TCATGCCAAG ATATAATAAA AGAAGAACTA TCTGTCTATA GGAAGGAGAC CGAGGAAGTG ACTGATGTAG AACTGTGTTT GACATTTTCA 3400
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 GACAAAACAT ACTGTACAAA CTTAAACCCT GTTACCCCTA TTATAACAGA TCTATTTGAG GTACAGTTCA AAACTGTAGA GACCTACAGC TTGCCCAGAA 3500
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 TTGTTGCTGT GCAAAACCAT GAGATTAAGA TTGGGCAAAT AAATGATTTA GGAGTTTACT CTAAGGGTTG TGGGAATGTT CAAAAGGTCA ATGGAACTAT 3600
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007

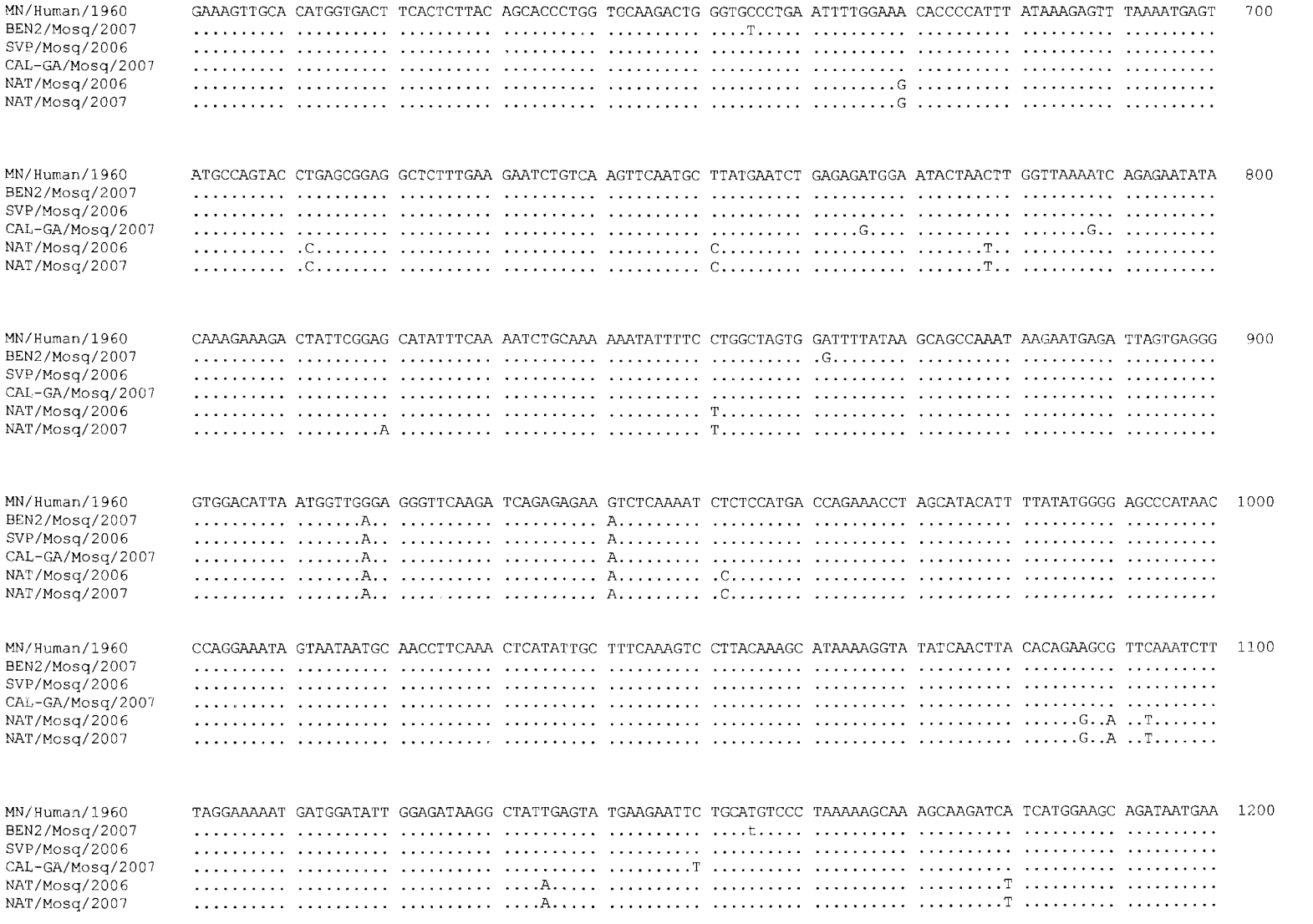
```
MN/Human/1960     TTATGGCAAT GGAGTTCCCA GATTTGACTA CTTATGCCAT TTAGCTAGCA GGAAGGAAGT CATTGTTAGA AAATGCTTCG ACAATGATTA CCAAGCATGC   3700
BEN2/Mosq/2007    .......... .......... A......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......T.. .........C .......... .A........ .......... .......... .......... ..........
NAT/Mosq/2007     .......... .......... .......T.. .........C .......... .A........ .......... .......... .......... ..........

MN/Human/1960     AAATTCCTTC AAAGCCCTGC TAGTTACAGA CTTGAAGAAG ACAGTGGCAC TGTGACCATA ATTGACTACA AAAAGATTTT AGGAACAATC AAGATGAAGG   3800
BEN2/Mosq/2007    .....T.... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .....T.... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .....T.... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......... .......... .......T.. ...A..T... .....T.... .G.....A.. ...C...... ..........
NAT/Mosq/2007     .......... .......... .......... .......... .......T.. ...A..T... .....T.... .G.....A.. ...C...... ..........

MN/Human/1960     CAATTTTAGG AGATGTCAAA TATAAAACAT TTGCTGATAG TGTCGATATA ACCGCAGAAG GGTCATGCAC CGGCTGTATT AACTGCTTCG AAAATATCCA   3900
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... ........G. .......... .......... ..T....... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......... .......C.. .......... ..T....... .......... T........C ..T.....T. .G.....T..
NAT/Mosq/2007     .......... .......... .......... .......C.. .......... ..T....... .......... T........C ..T.....T. .G.....T..

MN/Human/1960     TTGCGAATTA ACGTTGCACA CCACAATTGA AGCCAGCTGC CCAATTAAAA GCTCGTGCAC AGTATTTCAT GACAGGATTC TTGTGACTCC AAATGAACAC   4000
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     ...T...... .....A..T. .......... ...T...... .......... .......... ......C... .......... .......... ..........
NAT/Mosq/2007     ...T...... .....A..T. .......... ...T...... .......... .......... ......C... .......... .......... ..........

MN/Human/1960     AAATATGCAT TGAAAATGGT GTGCACAGAA AAGCCAGGGA ACACACTCAC AATTAAAGTC TGCAATACTA AAATTGAAGC ATCTATGGCC CTTGTCGATG   4100
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... .....A..C.
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .......... ........C.
CAL-GA/Mosq/2007  .......... .......... .......... .......... .......... .......... .......... .......... .......... ........C.
NAT/Mosq/2006     .......... .A........ .......... .G........ .T........ .........T .....C.... .G.....G.? .......... ..........
NAT/Mosq/2007     .......... .A........ .......... .G......g. .T........ .........T .....C.... .G.....G.. .......... ..........

MN/Human/1960     CAAAGCCTAT CATAGAACTA GCACCAGTTG ATCAGACAGC ATATATAAGA GAAAAAGATG AAAGGTGTAA AACTTGGATG TGTAGGGTAA GAGATGAAGG   4200
BEN2/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006     .......... .......... .......... .......... .......... .......... .......... .......... .....A.... ..........
CAL-GA/Mosq/2007  .......... .......T.. .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006     .......... .......... .......... .......... ...C...... .....G.... .......... .......... .......... ..........
NAT/Mosq/2007     .......... .......... .......... .......... ...C...... .....G.... .......... .......... .......... ..........
```

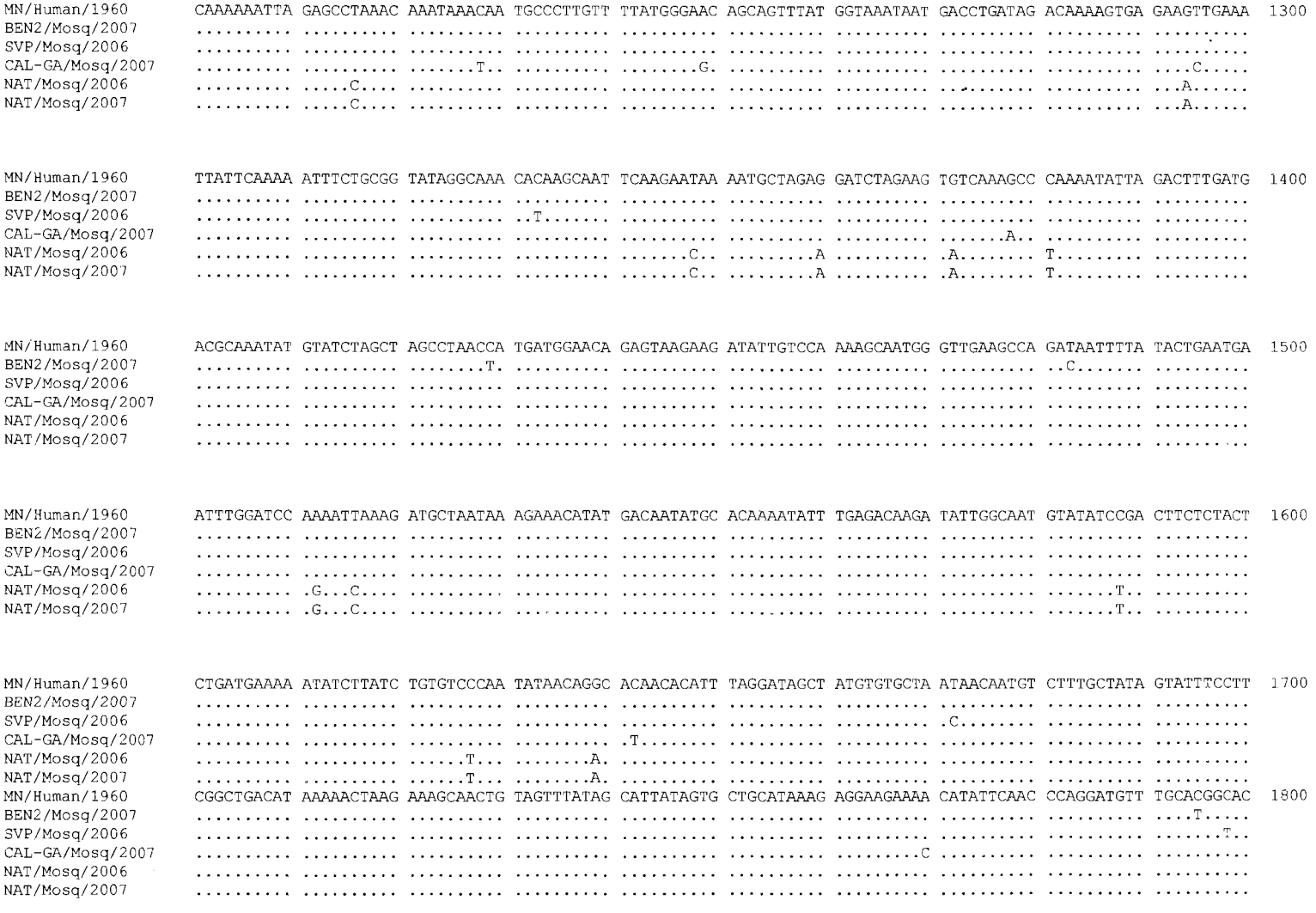
```
MN/Human/1960      ACTGCAGGTC ATCTTGGAGC CATTTAAAAA TTTATTTGGA TCTTATATTG GGATATTTTA CACATTTATT ATATCTATAG TAGCATTATT GGTTATTATC   4300
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      .T....A... ...C...... ....C..... ...G...... .......... .......... .......... .......... .G........ ......C...
NAT/Mosq/2007      .T....A... ...C...... ....C..... ...G...... .......... .......... .......... .......... .G........ ......C...


MN/Human/1960      TATGTACTAC TACCTATATG CTTTAAGTTA AGGGATACCC TTAGAAAGCA TGAAGATGCA TATAAGAGAG AGATGAAAAT TAGATAGGGG ATCTATGCAG   4400
BEN2/Mosq/2007     .......... ....C..... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .....C.... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... ..A....... .......... .......... .......... .......... .......... .......G..
NAT/Mosq/2006      .....G...T .G........ ...C...... ..A....... .......... .......... .......... .......... .......... G.ACG..T..
NAT/Mosq/2007      .....G...T .G........ ...C...... ..A....... .......... .......... .......... .......... .......... G.ACG..T..


MN/Human/1960      AACAAAATTG AGTCCTGTAT TATATATTC TATTTGTAGT ATAGCTGTTG TTAAGTGGGG GGTGGGGAAC TAACAACAGC GTAAATTTAT TTTGCAAACA   4500
BEN2/Mosq/2007     .......... .......... .......... .......... .......... .......... ..C....... .......... .......... ..........
SVP/Mosq/2006      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007   .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006      .......... ....T.A... .G....... C......... .......... .......... .......... .......... ..G....... ..........
NAT/Mosq/2007      .......... ....T.A..C .G....... C......... .......... .......... .......... .......... ..G....... ..........


MN/Human/1960      TTATTTTATA CTTGGTAGCA CACTACT    4527
BEN2/Mosq/2007     .......... .......... .......
SVP/Mosq/2006      .......... .......... .......
CAL-GA/Mosq/2007   .......... .......... .......
NAT/Mosq/2006      .......... .......... .......
NAT/Mosq/2007      .......... .......... .......
```

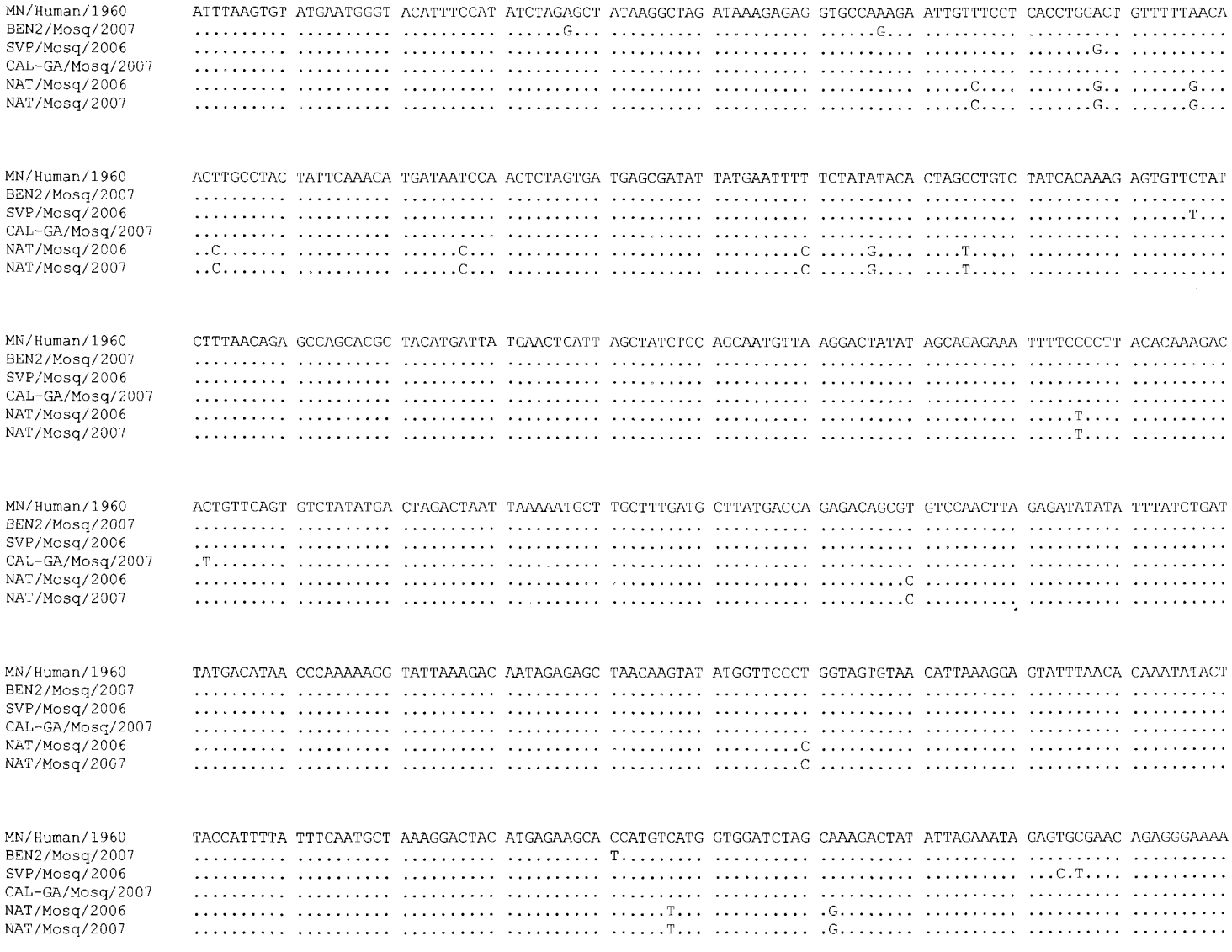
# APPENDIX 3

# Appendix 3: L segment sequence alignment of LACV prototype and viruses isolated from super-infected mosquitoes

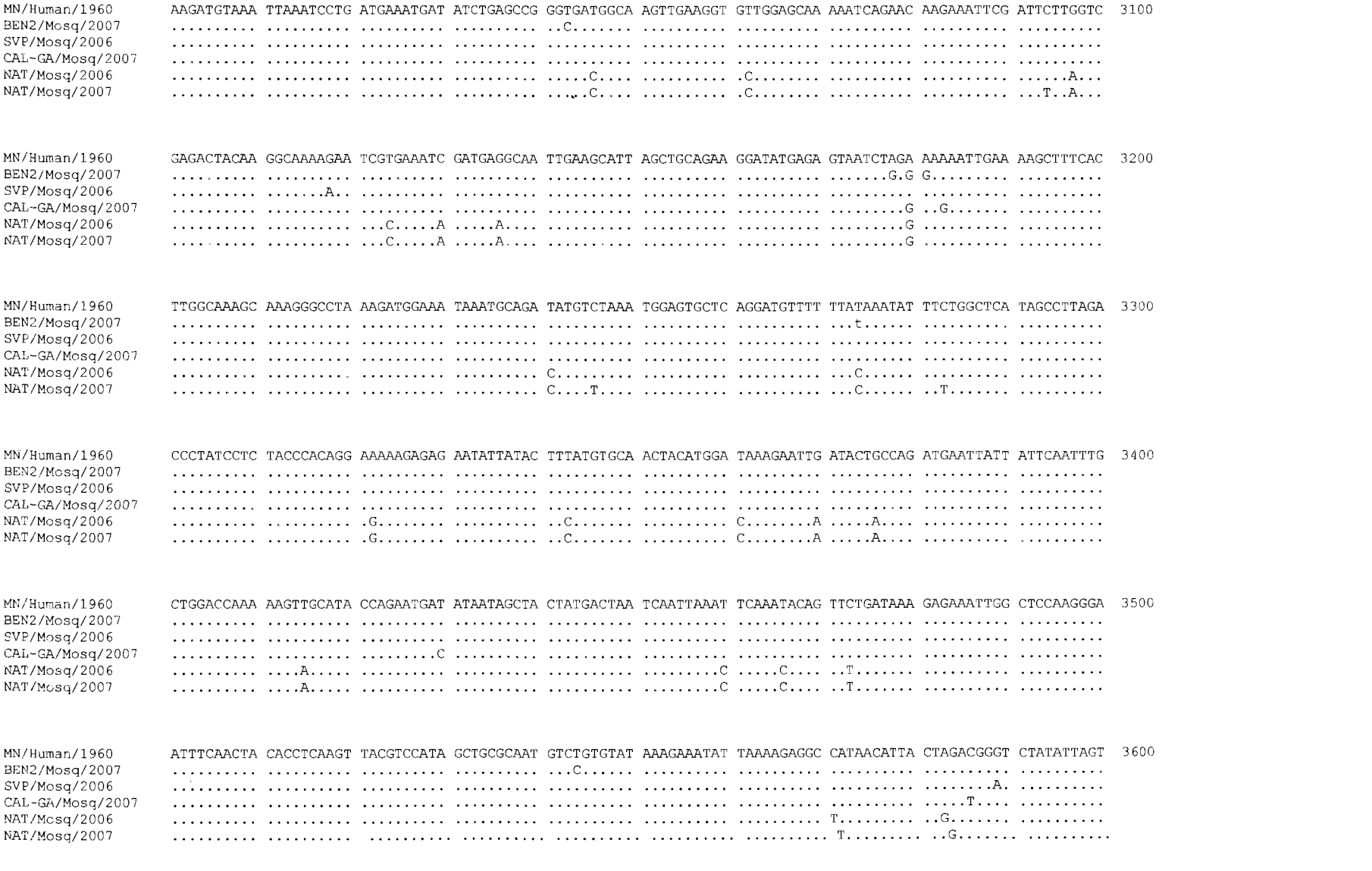
```
MN/Human/1960       AGTAGTGTAC TCCTATCTAC AAAACTTACA GAAAATTCAG TCATATCACA ATATATGCAT AATGGACTAT CAAGAGTATC AACAATTCTT GGCTAGGATT   100
BEN2/Mosq/2007      T..T..... C......... .......... .......... .......... .......... .......... .......... .......... A.........
SVP/Mosq/2006       T..T..... C......... .......... .......... .......... .......... .......... .......... .......... .........
CAL-GA/Mosq/2007    T..T..... C......... .......... .......... .......... .......... .......... .......... .......... .........
NAT/Mosq/2006       .......... C......... .......... A......... .......... .......... .........C .......... .......... .........C
NAT/Mosq/2007       T......... C......... .......... A......... .......... .......... .........C .......... .......... .........C

MN/Human/1960       AATACTGCAA GGGATGCATG TGTAGCCAAG GATATCGATG TTGACCTATT AATGGCCAGA CATGATTATT TTGGAAGAGA GCTGTGCAAG TCCTTAAATA   200
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... ....G..... ...A...... ..........
NAT/Mosq/2006       .......... ....C..... .......... .......... .C........ G......... .......... .......... ...A...... ..........
NAT/Mosq/2007       .......... ....C..... .......... .......... .C........ G......... .......... .......... ...A...... ..........

MN/Human/1960       TAGAATATAG GAATGATGTA CCATTTATAG ATATAATTTT GGATATAAGG CCCGAAGTAG ACCCATTAAC CATAGATGCA CCACATATTA CCCCAGACAA   300
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... A......... .......... .......... A......... ..T....... .......... .......... .......... ..........
NAT/Mosq/2007       .......... A......... .......... .......... A......... ..T....... .......... .......... .......... ..........

MN/Human/1960       TTATCTATAT ATAAATAATG TGTTATATAT CATAGATTAT AAGGTCTCTG TATCGAATGA AAGCAGTGTT ATAACATATG ACAAATATTA TGAGTTAACT   400
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       ....T..... .......... .......... .......... .......... .......... .......... .......... .......... ......G...
NAT/Mosq/2007       ....T..... .......... .......... .......... .......... .......... .......... .......... .......... ......G...

MN/Human/1960       AGGGACATAT CCGATAGATT AAGTATTCCA ATAGAAATAG TTATCGTCCG TATAGACCCT GTAAGTAAGG ATTTGCATAT TAACTCTGAT AGATTTAAAG   500
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... ........A. .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .....T.... .......... .......... .......... .......... .......... ........A. .......... .......... ..........
NAT/Mosq/2007       .....T.... .......... .......... .......... .......... .......... ........A. .......... .......... ..........

MN/Human/1960       AACTTTACCC TACAATAGTG GTGGATATAA ACTTCAATCA ATTTTTTGAC TTAAAACAAT TACTCTATGA AAAATTCGGT GATGATGAAG AATTCCTATT   600
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... ....T..... .......... .......... ..........
NAT/Mosq/2006       ....C..... .......... .......... .......... G......... .......... .......... .......... ..C....... ..........
NAT/Mosq/2007       ....C..... .......... .......... .......... G......... .......... .......... .......... ..C....... ..........
```


MN/Human/1960 GAAAGTTGCA CATGGTGACT TCACTCTTAC AGCACCCTGG TCCAAGACTG GGTGCCCTGA ATTTTGGAAA CACCCCATTT ATAAAGAGTT TAAAATGAGT 700
BEN2/Mosq/2007 ....T..... G
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006 G
NAT/Mosq/2007 G
MN/Human/1960 ATGCCAGTAC CTGAGCGGAG GCTCTTTGAA GAATCTGTCA AGTTCAATGC TTATGAATCT GAGAGATGGA ATACTAACTT GGTTAAAATC AGAGAATATA 800
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007 .....G.... .....G..
NAT/Mosq/2006 .C..... C..... .....T..
NAT/Mosq/2007 .C..... C..... .....T..
MN/Human/1960 CAAAGAAAGA CTATTCGGAG CATATTTCAA AATCTGCAAA AAATATTTTC CTGGCTAGTG GATTTTATAA GCAGCCAAAT AAGAATGAGA TTAGTGAGGG 900
BEN2/Mosq/2007 .G.....
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006 T.....
NAT/Mosq/2007 .....A T.....
MN/Human/1960 GTGGACATTA ATGGTTGGGA GGGTTCAAGA TCAGAGAGAA GTCTCAAAAT CTCTCCATGA CCAGAAACCT AGCATACATT TTATATGGGG AGCCCATAAC 1000
BEN2/Mosq/2007 .....A.. A.....
SVP/Mosq/2006 .....A.. A.....
CAL-GA/Mosq/2007 .....A.. A.....
NAT/Mosq/2006 .....A.. A..... .C.....
NAT/Mosq/2007 .....A.. A..... .C.....
MN/Human/1960 CCAGGAAATA GTAATAATGC AACCTTCAAA CTCATATTGC TTTCAAAGTC CTTACAAAGC ATAAAAGGTA TATCAACTTA CACAGAAGCG TTCAAATCTT 1100
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006 ......G..A ..T.....
NAT/Mosq/2007 ......G..A ..T.....
MN/Human/1960 TAGGAAAAAT GATGGATATT GGAGATAAGG CTATTGAGTA TGAAGAATTC TGCATGTCCC TAAAAAGCAA AGCAAGATCA TCATGGAAGC AGATAATGAA 1200
BEN2/Mosq/2007 ....t.....
SVP/Mosq/2006
CAL-GA/Mosq/2007 .....T
NAT/Mosq/2006 ....A..... .....T
NAT/Mosq/2007 ....A..... .....T


MN/Human/1960 CAAAAAATTA GAGCCTAAAC AAATAAACAA TGCCCTTGTT TTATGGGAAC AGCAGTTTAT GGTAAATAAT GACCTGATAG ACAAAAGTGA GAAGTTGAAA 1300
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... .....T.. ..... ........G. ..... ..... ..... ....C.....
NAT/Mosq/2006 ..... .....C.... ..... ..... ..... ..... ..... ..... ...A.....
NAT/Mosq/2007 ..... .....C.... ..... ..... ..... ..... ..... ..... ...A.....
MN/Human/1960 TTATTCAAAA ATTTCTGCGG TATAGGCAAA CACAAGCAAT TCAAGAATAA AATGCTAGAG GATCTAGAAG TGTCAAAGCC CAAAATATTA GACTTTGATG 1400
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..T..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... .....A.. ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... .....C.. .....A ..... .A..... T..... .....
NAT/Mosq/2007 ..... ..... ..... ..... .....C.. .....A ..... .A..... T..... .....
MN/Human/1960 ACGCAAATAT GTATCTAGCT AGCCTAACCA TGATGGAACA GAGTAAGAAG ATATTGTCCA AAAGCAATGG GTTGAAGCCA GATAATTTTA TACTGAATGA 1500
BEN2/Mosq/2007 ..... ..... .....T. ..... ..... ..... ..... ..... ..C..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
MN/Human/1960 ATTTGGATCC AAAATTAAAG ATGCTAATAA AGAAACATAT GACAATATGC ACAAAATATT TGAGACAAGA TATTGGCAAT GTATATCCGA CTTCTCTACT 1600
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 .....G...C.... ..... ..... ..... ..... ..... ..... .....T.. .....
NAT/Mosq/2007 .....G...C.... ..... ..... ..... ..... ..... ..... .....T.. .....
MN/Human/1960 CTGATGAAAA ATATCTTATC TGTGTCCCAA TATAACAGGC ACAACACATT TAGGATAGCT ATGTGTGCTA ATAACAATGT CTTTGCTATA GTATTTCCTT 1700
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... .C..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... .T..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... .....T... .....A. ..... ..... ..... ..... .....
NAT/Mosq/2007 ..... ..... .....T... .....A. ..... ..... ..... ..... .....
MN/Human/1960 CGGCTGACAT AAAAACTAAG AAAGCAACTG TAGTTTATAG CATTATAGTG CTGCATAAAG AGGAAGAAAA CATATTCAAC CCAGGATGTT TGCACGGCAC 1800
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ....T.....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... .....T..
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... .....C ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....


MN/Human/1960 ATTTAAGTGT ATGAATGGGT ACATTTCCAT ATCTAGAGCT ATAAGGCTAG ATAAAGAGAG GTGCCAAAGA ATTGTTTCCT CACCTGGACT GTTTTTAACA 1900
BEN2/Mosq/2007 ..... ..... ..... .....G... ..... ..... .....G... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... .....G.. .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... .....C.... .....G.. .....G...
NAT/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... .....C.... .....G.. .....G...
MN/Human/1960 ACTTGCCTAC TATTCAAACA TGATAATCCA ACTCTAGTGA TGAGCGATAT TATGAATTTT TCTATATACA CTAGCCTGTC TATCACAAAG AGTGTTCTAT 2000
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....T...
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..C..... ..... .....C... ..... ..... .....C .....G.... ....T..... ..... .....
NAT/Mosq/2007 ..C..... ..... .....C... ..... ..... .....C .....G.... ....T..... ..... .....
MN/Human/1960 CTTTAACAGA GCCAGCACGC TACATGATTA TGAACTCATT AGCTATCTCC AGCAATGTTA AGGACTATAT AGCAGAGAAA TTTTCCCCTT ACACAAAGAC 2100
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... .....T.... .....
NAT/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... .....T.... .....
MN/Human/1960 ACTGTTCAGT GTCTATATGA CTAGACTAAT TAAAAATGCT TGCTTTGATG CTTATGACCA GAGACAGCGT GTCCAACTTA GAGATATATA TTTATCTGAT 2200
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 .T..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... ..... .....C ..... ..... .....
NAT/Mosq/2007 ..... ..... ..... ..... ..... ..... .....C ..... ..... .....
MN/Human/1960 TATGACATAA CCCAAAAAGG TATTAAAGAC AATAGAGAGC TAACAAGTAT ATGGTTCCCT GGTAGTGTAA CATTAAAGGA GTATTTAACA CAAATATACT 2300
BEN2/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ..... .....C ..... ..... ..... .....
NAT/Mosq/2007 ..... ..... ..... ..... ..... .....C ..... ..... ..... .....
MN/Human/1960 TACCATTTTA TTTCAATGCT AAAGGACTAC ATGAGAAGCA CCATGTCATG GTGGATCTAG CAAAGACTAT ATTAGAAATA GAGTGCGAAC AGAGGGAAAA 2400
BEN2/Mosq/2007 ..... ..... ..... ..... T..... ..... ..... ..... ..... .....
SVP/Mosq/2006 ..... ..... ..... ..... ..... ..... ..... ..... ...C.T.... .....
CAL-GA/Mosq/2007 ..... ..... ..... ..... ..... ..... ..... ..... ..... .....
NAT/Mosq/2006 ..... ..... ..... ..... ......T... ..... .G..... ..... ..... .....
NAT/Mosq/2007 ..... ..... ..... ..... ......T... ..... .G..... ..... ..... .....

```
MN/Human/1960       CATAAAGGAG ATATGGTCTA CAAATTGTAC CAAACAGACA GTGAACCTTA AAATTTTGAT CCATTCCTTG TGCAAGAATT TACTAGCAGA CACTTCAAGA  2500
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... A......... .......... .......... .......... .......... .G........ ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... ....G..... ..........
NAT/Mosq/2007       .......... .......... .......... .......... .......... .......... .......... .......... ....G..... ..........

MN/Human/1960       CACAACCACT TGCGGAACAG AATAGAAAAT AGGAACAATT TTAGAAGGTC TATAACAACT ATTTCAACAT TTACAAGTTC GAAGTCTTGC CTCAAAATAG  2600
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .....G.... .......... .......... ..........
SVP/Mosq/2006       .......... ... ...... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... ......G... .......... .......... .......... .....G.... .......... .......... ..........
NAT/Mosq/2007       .......... .......... ......G... .......... .......... .......... .....G.... .......... .......... ..........

MN/Human/1960       GGGACTTTAG AAAAGAGAAA GAGCTGCAGT CAGTTAAACA GAAGAAAATC TTAGAGGTGC AGAGTCGCAA AATGAGATTA GCAAACCCAA TGTTCGTGAC  2700
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .....T.... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ....T.....
NAT/Mosq/2006       .......... ......A... .......... .......... .......... .......... .......... .......... .....T.... .......A..
NAT/Mosq/2007       .......... ......A... .......... .......... .......... .......... .......... .......... .....T.... .......A..

MN/Human/1960       AGATGAACAA GTATGCCTTG AAGTTGGGCA CTGCAATTAT GAGATGCTGA GGAATGCTAT GCCGAATTAT ACAGATTATA TATCAACTAA AGTATTTGAT  2800
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... ..G....... .G........ .......... .......... .......... .......... .......... .......... ......????
NAT/Mosq/2007       .......... ..G....... .G........ .......... .......... .......... .......... .......... .......... ..........

MN/Human/1960       AGGTTATATG AGTTATTAGA TAAAAGAGTT TTGACAGACA AGCCTGTTAT AGAGCAAATA ATGGATATGA TGGTCGACCA CAAAAAGTTC TATTTCACAT  2900
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... ...G...... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ???....... ..........
NAT/Mosq/2007       .......... ....G..... .....A.... ..A....... .......... .......... .......... .......... .......... ..........

MN/Human/1960       TTTTCAATAA AGGCCAGAAA ACGTCAAAGG ATAGAGAGAT ATTCGTTGGA GAATATGAAG CTAAAATGTG TATGTACGCA GTTGAGAGAA TAGCAAAAGA  3000
BEN2/Mosq/2007      .......... .......... .......... .......... ...T...... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... ..A....... .......... .......... .......... .......... .......... ........G. ..........
NAT/Mosq/2007       .......... .......... ..A....... .......... .......... .......... .......... .......... ........G. ..........
```

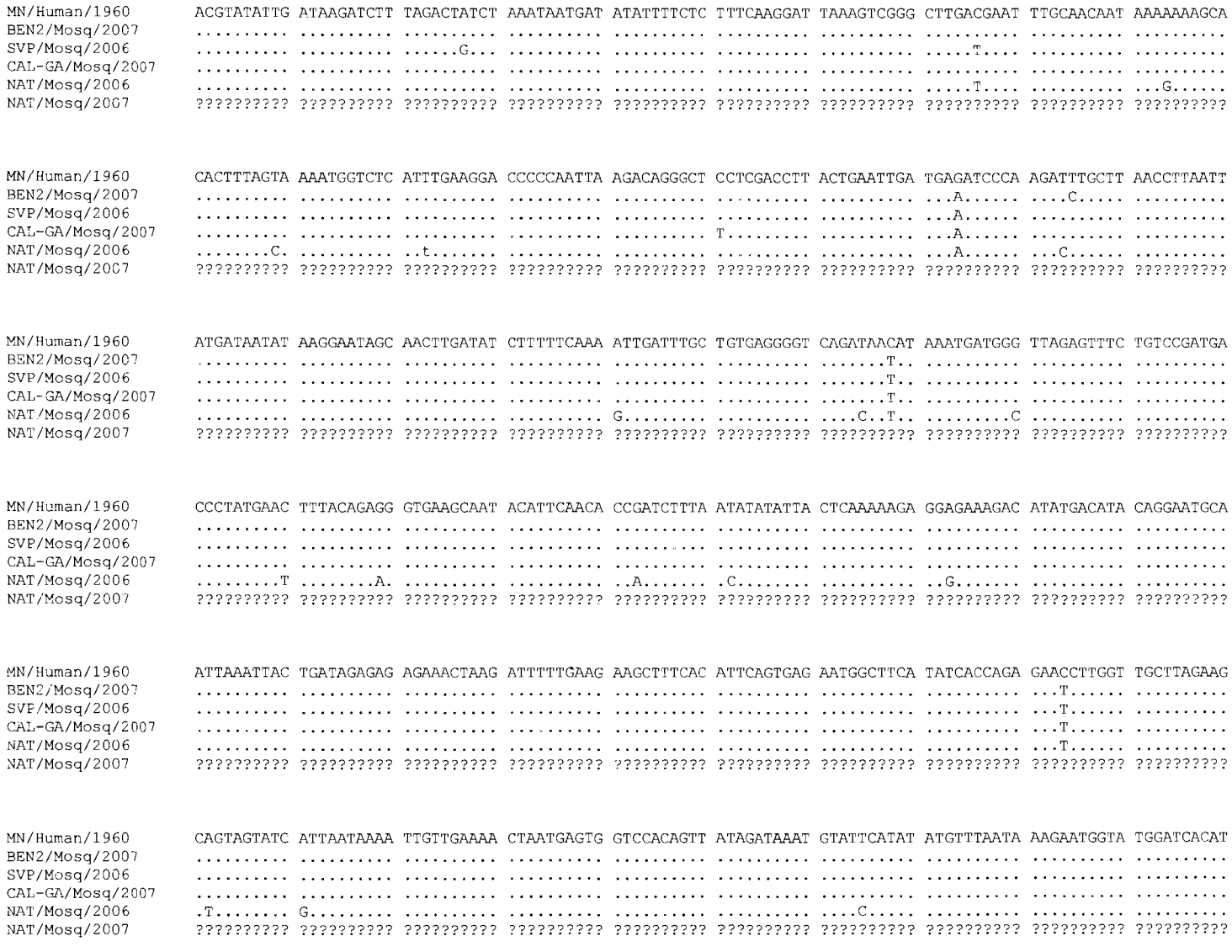

MN/Human/1960 AAGATGTAAA TTAAATCCTG ATGAAATGAT ATCTGAGCCG GGTGATGGCA AGTTGAAGGT GTTGGAGCAA AAATCAGAAC AAGAAATTCG ATTCTTGGTC 3100
BEN2/Mosq/2007 ..C.....
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006 .....C.... .....C.... .....A...
NAT/Mosq/2007 .....C.... .....C.... ...T..A...
MN/Human/1960 GAGACTACAA GGCAAAAGAA TCGTGAAATC GATGAGGCAA TTGAAGCATT AGCTGCAGAA GGATATGAGA GTAATCTAGA AAAAATTGAA AAGCTTTCAC 3200
BEN2/Mosq/2007 .....G.G G.....
SVP/Mosq/2006 .....A..
CAL-GA/Mosq/2007 .....G ..G.....
NAT/Mosq/2006 ...C.....A .....A..... .....G
NAT/Mosq/2007 ...C.....A .....A..... .....G
MN/Human/1960 TTGGCAAAGC AAAGGGCCTA AAGATGGAAA TAAATGCAGA TATGTCTAAA TGGAGTGCTC AGGATGTTTT TTATAAATAT TTCTGGCTCA TAGCCTTAGA 3300
BEN2/Mosq/2007 ...t.....
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006 C.....C.....
NAT/Mosq/2007 C....T..... ...C..... ..T.....
MN/Human/1960 CCCTATCCTC TACCCACAGG AAAAAGAGAG AATATTATAC TTTATGTGCA ACTACATGGA TAAAGAATTG ATACTGCCAG ATGAATTATT ATTCAATTTG 3400
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006 .G..... .....C..... C.....A .....A.....
NAT/Mosq/2007 .G..... .....C..... C.....A .....A.....
MN/Human/1960 CTGGACCAAA AAGTTGCATA CCAGAATGAT ATAATAGCTA CTATGACTAA TCAATTAAAT TCAAATACAG TTCTGATAAA GAGAAATTGG CTCCAAGGGA 3500
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007 .....C
NAT/Mosq/2006 ....A..... .....C..... ..T.....
NAT/Mosq/2007 ....A..... .....C..... ..T.....
MN/Human/1960 ATTTCAACTA CACCTCAAGT TACGTCCATA GCTGCGCAAT GTCTGTGTAT AAAGAAATAT TAAAAGAGGC CATAACATTA CTAGACGGGT CTATATTAGT 3600
BEN2/Mosq/2007 ...C.....
SVP/Mosq/2006 .....A.
CAL-GA/Mosq/2007 .....T.....
NAT/Mosq/2006 T..... ..G.....
NAT/Mosq/2007 T..... ..G.....

```
MN/Human/1960       CAACTCATTA GTCCATTCGG ATGATAACCA AACATCGATA ACAATAGTTC AGGATAAGAT GGAAAATGAT AAAATTATAG ATTTTGCAAT GAAAGAATTT  3700
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... ..T....... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       ...T...... ........T. ....C..... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2007       ...T...... ........T. ....C..... .......... .......... .......... .......... .......... .......... ..........

MN/Human/1960       GAGAGAGCCT GTTTGACATT TGGATGCCAA GCAAATATGA AAAAGACATA TGTAACAAAT TGTATAAAAG AGTTTGTTTC ATTATTTAAC TTGTATGGCG  3800
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......C.. .......... .....C....
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... ...G...... .......... .......... .......... .....C....
NAT/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......C.. .........T ........T.
NAT/Mosq/2007       .......... .......... .......... .......... .......... .......... .......... .......C.. .........T ........T.

MN/Human/1960       AACCCTTTTC AATATATGGC AGATTCCTAT TAACATCTGT GGGTGATTGT GCCTATATAG GGCCTTATGA AGATTTAGCT AGTCGAATAT CATCAGCCCA  3900
BEN2/Mosq/2007      .......... .......... ..G....... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       ....T..... .......... .......... .G........ .......... .......... .......... ......G... .......... ..........
NAT/Mosq/2007       ....T..... .......... .......... .G........ .......... .......... .......... ......G... .......... ..........

MN/Human/1960       GACAGCCATA AAGCATGGTT GTCCACCCAG TCTAGCATGG GTGTCCATAG CAATAAGTCA TTGGATGACC TCTCTGACAT ACAACATGCT ACCAGGGCAG  4000
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .....G.... .......C.. .......... .....A.... ........T. ..........
NAT/Mosq/2007       .......... .......... .......... .......... .....G.... .......C.. .......... .....A.... ........T. ..........

MN/Human/1960       TCAAATGACC CAATTGATTA TTTCCCTGCA GAAAATAGGA AGGATATCCC TATAGAATTG AATGGTGTGT TAGATGCTCC GTTGTCAATG ATTAGTACAG  4100
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... ........A. .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... ........A. ....C..... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... ........A. .......... .......... ..........
NAT/Mosq/2006       .......... .......... ......A... ..G....... .......... C......... ........A. .......... .......... ..........
NAT/Mosq/2007       .......... .......... ......A... ..G....... .......... C......... ........A. .......... .......... ..........

MN/Human/1960       TTGGATTGGA ATCTGGGAAT TTATACTTCT TGATAAAGTT GTTGAGCAAA TACACCCCGG TCATGCAGAA AAGAGAGTCA GTAGTCAACC AAATAGCTGA  4200
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... ..T....... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .........C .......... .......A.. ......T... .......... .......... .......... .......... ..........
NAT/Mosq/2007       .......... .........C .......... .......A.. ......T... .......... .......... .......... .......... ..........
```

```
MN/Human/1960       AGTTAAGAAC TGGAAGGTCG AGGATCTAAC AGATAATGAA ATATTTAGGC TTAAAATACT CAGATATTTA GTTCTAGATG CAGAGATGGA CCCTAGTGAT  4300
BEN2/Mosq/2007      .......... .......... .......... .......... ........A. .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... ........A. .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... ..........A......... .......... ........A. .......... .......... .......... .......... ..........
NAT/Mosq/2006       ...CC..... .......... .......... .......... ........A. .......... .......... .......... .......... ..........
NAT/Mosq/2007       .......... .......... .......... .......... ........A. .......... .......... .......... .......... ..........

MN/Human/1960       ATTATGGGTG AGACAAGCGA CATGAGAGGG AGGTCCATTT TGACACCTAG AAAATTCACA ACAGCAGGCA GTTTAAGGAA ATTATATTCT TTCAGTAAGT  4400
BEN2/Mosq/2007      .......... .......... .......... .....T.... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .....T.... .......... .......... .......... .......... ......C... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .....T.... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... ..A..T.... .A........ .......... ........T. .......... .......... ..........
NAT/Mosq/2007       .......... .......... .......... ..A..T.... .A........ .......... ........T. .......... .......... ..........

MN/Human/1960       ACCAGGATAG ACTGTCTTCC CCTGGAGGCA TGGTTGAATT GTTCACTTAT TTGCTTGAGA AACCTGAGTT GTTAGTGACT AAAGGAGAAG ATATGAAAGA  4500
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ......G...
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... ..???????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
NAT/Mosq/2007       .......... .......... ..C....... .......... .......... .......... .......... ..........       .......... ..........

MN/Human/1960       TTATATGGAA TCTGTGATAT TCCGATATAA CTCCAAAAGG TTCAAAGAAA GTTTGTCAAT ACAGAACCCA GCACAATTAT TTATAGAACA GATATTGTTC  4600
BEN2/Mosq/2007      .......... .......... .......... T......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... T......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... T......... ..T....... ..C....... .......... .......... .......... ..........
NAT/Mosq/2006       ?????????.. .......... .......C.. T......... .......... .......... ......T... .......... .......... ..........
NAT/Mosq/2007       .......... .......... .......C.. T......... .......... .......... ......T... .......... .......... ..........

MN/Human/1960       TCACATAAGC CCATAATAGA CTTTTCTGGT ATCAGGGACA AATATATAAA CCTACATGAT AGTAGAGCTC TAGAGAAGGA ACCTGACATA TTAGGAAAAG  4700
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... .....G....
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... ....A..... ......T... ..G.......
NAT/Mosq/2007       .......... .......... .......... .......... .......... .......... .......... ....A..... ......T... ..G.......

MN/Human/1960       TAACATTTAC AGAGGCTTAT AGATTATTAA TGAGGGACCT GTCTAGCCTA GAACTAACCA ATGATGACAT CCAAGTAATT TATTCTTACA TAATACTTAA  4800
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......T.. A......... .......... .......... .......... .......... ..........
NAT/Mosq/2007       .......... .......... .......... .......T.. A......... .......... .......... .......... .......... ..........
```

```
MN/Human/1960       TGACCCTATG ATGATAACTA TTGCAAACAC ACATATATTG TCAATATACG GGAGTCCTCA ACGGAGGATG GGCATGTCCT GTTCAACGAT GCCAGAATTT  4900
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .......... .A..C..... .......... ..T....... .......... ..........
NAT/Mosq/2007       .......... .......... .......... .......... .......... .A..C..... .......... ..T....... .......... ..........

MN/Human/1960       AGAAATTTAA AATTAATACA TCATTCCCCA GCCTTAGTTT TGAGAGCATA TAGTAAAAAT AATCCTGACA TCCAGGGTGC TGATCCCACG GAAATGGCTA  5000
BEN2/Mosq/2007      .....C.... .......... .......... .......... .......... ...C...... ..C....... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .T........ .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       ..G..C.... .......... .......... ..T....... .......... .......... .......... .......... .........A ..........
NAT/Mosq/2007       ..G..C.... .......... .......... ..T....... .......... .......... .......... .......... .........A ..........

MN/Human/1960       GAGATTTAGT TCATCTGAAA GAGTTCGTTG AGAATACAAA TTTAGAAGAA AAAATGAAAG TTAGGATTGC TATAAATGAA GCAGAGAAAG GACAACGGGA  5100
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... ..G....... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... ......A... .......... .......... ...G...... .......... .......... .......... .......... .......A..
NAT/Mosq/2007       .......... ......A... .......... .......... ...G...... .......... .......... .......... .......... .......A..

MN/Human/1960       TATAGTCTTT GAACTAAAAG AGATGACTAG ATTTTATCAG GTTTGCTATG AGTATGTCAA ATCTACAGAA CACAAGATAA AAGTCTTCAT TCTCCCGACA  5200
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......C. .......T.. .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... ...C...... .......... .......... .......... .......... .......... ...T...G..
NAT/Mosq/2007       .......... .......... .......... ...C...... .......... .......... .......... .......... .......... ...T...G..

MN/Human/1960       AAATCATACA CAACAACAGA TTTCTGTTCA CTCATGCAGG GGAATTTAAT AAAAGATAAA GAGTGGTACA CAGTTCACTA CCTAAAACAG ATATTGTCTG  5300
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .......... ...G...... .......... ....C..... ...G...... ..........
NAT/Mosq/2007       .......... .......... .......... .......... .......... ...G...... .......... ....C..... ...G...... ..........

MN/Human/1960       GTGGCCATAA AGCCATAATG CAGCATAATG CCACTAGTGA GCAAAATATT GCTTTTGAGT GTTTCAAATT AATTACCCAT TTTGCAGACT CATTCATAGA  5400
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... G......... .......... .......... .......... .......... ....T..... G......... .......... ..........
NAT/Mosq/2007       .......... G......... .......... .......... .......... .......... ....T..... G......... .......... ..........
```

```
MN/Human/1960       TTCATTATCT AGGTCAGCTT TTTTGCAGTT GATAATAGAT GAATTCAGTT ATAAAGATGT GAAGGTTAGC AAACTTTATG ACATAATAAA GAATGGGTAT  5500
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... A......... .......... .......... ..........
NAT/Mosq/2006       .......... ........C. .......... .......... .......... ....G..... ........C. .......... .......... ..........
NAT/Mosq/2007       .......... ........C. .......... .......... .......... ....G..... ........C. .......... .......... ..........

MN/Human/1960       AATCGAACTG ACTTCATACC ATTGCTTTTT AGAACTGGCG ATTTAAGACA AGCTGACTTA GACAAGTATG ATGCTATGAA AAGTCATGAG AGGGTTACAT  5600
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .........C .......... .......... G......... .....A.... ....C..... ......C... ..........
NAT/Mosq/2007       .......... ..........  .........C .......... .......... G......... .....A.... ....C..... ......C... ..........

MN/Human/1960       GGAATGATTG GCAAACATCT CGTCACTTGG ACATGGGTTC AATTAATCTA ACAATAACCG GTTACAATAG ATCAATAACA ATAATCGGAG AAGATAACAA  5700
BEN2/Mosq/2007      .......... .......... .......... .T........ .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......C.. .......... .......... .......... .......... .......... ....T..... G......... .......... ..........
NAT/Mosq/2007       .......C.. .......... .......... .......... .......... .......... ....T..... G......... .......... ..........

MN/Human/1960       ATTGACATAT GCAGAATTAT GTCTAACTAG GAAAACTCCC GAGAATATAA CTATAAGTGG CAGAAAATTG CTAGGTGCAA GGCATGGACT TAAATTTGAA  5800
BEN2/Mosq/2007      .......... .......... .......... ......C..T .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .........T .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .........T .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       G......... .....GC... .......... .........T .......... .......C.. ......G... ......T... .......G.. ..........
NAT/Mosq/2007       G......... .....GC... .......... .........T .......... .......C.. ......G... ......T... .......G.. ..........

MN/Human/1960       AATATGTCCA AAATCCAAAC ATACCCAGGC AATTATTATA TAACATATAG AAAGAAAGAT CGCCACCAGT TTGTATACCA GATACATTCT CATGAATCAA  5900
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .......... ......G... .......... .......... .......... ..........
NAT/Mosq/2007       .......... .......... .......... .......... .......... ......G... .......... .......... .......... ..........

MN/Human/1960       TAACAAGGAG GAATGAAGAG CATATGGCTA TCAGGACCAG AATATACAAT GAAATAACCC CAGTATGTGT AGTTAACGTT GCAGAGGTGG ATGGGGACCA  6000
BEN2/Mosq/2007      .......... .......... .......... .......... .......... ........T. .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... ........T. .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... A......... .......... .......... .......... ........T. .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... G......... ........T. .......... ......T... .......... ..........
NAT/Mosq/2007       .......... .......... .......... .......... G......... ........T. .......... .......??? ?????????? ??????????
```


MN/Human/1960 ACGTATATTG ATAAGATCTT TAGACTATCT AAATAATGAT ATATTTTCTC TTTCAAGGAT TAAAGTCGGG CTTGACGAAT TTGCAACAAT AAAAAAAGCA 6100
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 CACTTTAGTA AAATGGTCTC ATTTGAAGGA CCCCCAATTA AGACAGGGCT CCTCGACCTT ACTGAATTGA TGAGATCCCA AGATTTGCTT AACCTTAATT 6200
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 ATGATAATAT AAGGAATAGC AACTTGATAT CTTTTTCAAA ATTGATTTGC TGTGAGGGGT CAGATAACAT AAATGATGGG TTAGAGTTTC TGTCCGATGA 6300
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 CCCTATGAAC TTTACAGAGG GTGAAGCAAT ACATTCAACA CCGATCTTTA ATATATATTA CTCAAAAAGA GGAGAAAGAC ATATGACATA CAGGAATGCA 6400
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 ATTAAATTAC TGATAGAGAG AGAAACTAAG ATTTTTGAAG AAGCTTTCAC ATTCAGTGAG AATGGCTTCA TATCACCAGA GAACCTTGGT TGCTTAGAAG 6500
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007
MN/Human/1960 CAGTAGTATC ATTAATAAAA TTGTTGAAAA CTAATGAGTG GTCCACAGTT ATAGATAAAT GTATTCATAT ATGTTTAATA AAGAATGGTA TGGATCACAT 6600
BEN2/Mosq/2007
SVP/Mosq/2006
CAL-GA/Mosq/2007
NAT/Mosq/2006
NAT/Mosq/2007

```
MN/Human/1960       GTACCATTCA TTTGATGTCC CTAAATGTTT TATGGGGAAT CCTATCACTA GAGACATAAA TTGGATGATG TTTAGAGAAT TCATCAATAG TTTACCAGGG  6700
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .....A.... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2007       ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????

MN/Human/1960       ACAGATATAC CACCATGGAA TGTCATGACA GAGAACTTCA AAAAGAAATG TATTGCTCTG ATAAACTCTA AGTTAGAAAC ACAGAGAGAT TTCTCAGAAT  6800
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
SVP/Mosq/2006       .......... .......... .......... .......... ...G...... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... .......... ...T...... .......... .......... .G........ .......... ....G..... .......... ..T.......
NAT/Mosq/2007       ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????

MN/Human/1960       TCACTAAACT GATGAAAAAG GAAGGTGGGA GGAGTAATAT AGAGTTTGAT TAGTAGTTAT GAGTTTACAG AGAACCTACA ATTAGGCCAT AAATTTGGGA  6900
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .......... ......C... .......... ..........
SVP/Mosq/2006       .......G.. .......... .......... .......... .......... .......... .......... .......... .......... ..........
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .......... .......... .......... ..........
NAT/Mosq/2006       .......... ......G... .......... ....C..... .......... .......... A......... .....TG... .......... ..........
NAT/Mosq/2007       ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????

MN/Human/1960       GGGTTTTGGA AATTGGCTAA AATTCAAAAA GAGGGGGATT AACAGCAACT GTATAAATTT GTAGATAGGA GCACACTACT  6980
BEN2/Mosq/2007      .......... .......... .......... .......... .......... .......... .........G ..........
SVP/Mosq/2006       .......... .......... .......... .......... .......... .......... .........G .........A
CAL-GA/Mosq/2007    .......... .......... .......... .......... .......... .......... .........G ..........
NAT/Mosq/2006       .......... .......... G......... .......... .......... .......... .........G ....C.....
NAT/Mosq/2007       ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ?????????? ??????????
```

# REFERENCES

Aitken, T. H., Tesh, R., Beaty, B. J. & Rosen, L. (1979). Transovarial transmission of yellow fever virus by mosquitoes (*Aedes aegypti*). *American Journal of Tropical Medicine and Hygiene* **28**, 119-121.

Amundson, T. & Yuill, T. (1981). Natural La Crosse virus infection in the red fox (*Vulpes fulva*), gray fox (*Urocyon cinereoargenteus*), raccoon (*Procyn lotor*), and opossum (*Didelphis virginiana*). *American Journal of Tropical Medicine and Hygiene* **30**, 706-714.

Anthony, C. & Jallon, J. M. (1982). The chemical basis for sex recognition in *Drosophila melanogaster*. *Journal of Insect Physiology* **28**, 873-880.

Armstrong, P. M. & Andreadis, T. G. (2006). A new genetic variant of La Crosse virus (*Bunyaviridae*) isolated from New England. *American Journal of Tropical Medicine and Hygiene* **75**, 491-496.

Armstrong, P. M. & Andreadis, T. G. (2007). Genetic relationships of Jamestown Canyon virus strains infecting mosquitoes collected in Connecticut. *American Journal of Tropical Medicine and Hygiene* **77**, 1157-1162.

Avsic-Zupanc, T., Petrovec, M., Duh, D., Lundkvist, A. & Plyusnin, A. (2007). Puumala hantavirus in Slovenia: Analyses of S and M segment sequences recovered from patients and rodents. *Virus Research* **123**, 204-210.

Baldridge, G. D., Beaty, B. J. & Hewlett, M. J. (1989). Genomic stability of La Crosse virus during vertical and horizontal transmission. *Archives of Virology* **108**, 89-99.

Balfour, H. H., Siem, R. A., Bauer, H. & Quie, P. G. (1973). California encephalitis (La Crosse) infection: I Clinical and laboratory findings in 66 children with meningoencephalitis. *Pediatrics* **52**, 680.

Bardos, V. (1965). Tahyna virus in experimentally infected white mice. *Acta Virology* **9**, 358-366.

Barrett, A. D., Monath, T. P., Cropp, C. B., Adkins, J. A., Ledger, T. N., Gould, E., Schlesinger, J. J., Kinney, R. M. & Trent, D. W. (1990). Attenuation of wild-type yellow fever virus by passage in HeLa cells. *Journal of General Virology* **71**, 2301-2306.

Beaty, B. J. & Bishop, D. H. L. (1988). Bunyavirus-vector interactions. *Virus Research* **10**, 289-300.

Beaty, B. J. & Calisher, C. H. (1991). Bunyaviridae - Natural-History. *Current Topics in Microbiology and Immunology* **169**, 27-78.

Beaty, B. J., Fuller, F. & Bishop, D. H. L. (1983). Bunyavirus gene structure: function relationships and potential for RNA segment reassortment in the vector: La Crosse and Snowshoe hare reassortant viruses in mosquitoes. In *California serogroup viruses*, pp. 119-128. Edited by C. H. Calisher & W. H. Thompson. N.Y.: A.R. Liss Inc.

Beaty, B. J., Holterman, M., Tabachnick, W., Shope, R. E., Rozhon, E. J. B. & Bishop, D. H. L. (1981). Molecular basis of bunyavirus transmission by mosquitoes: Role of the middle-sized RNA segment. *Science* **211**, 1433-1434.

Beaty, B. J., Miller, B. R., Shope, R. E., Rozhon, E. J. & Bishop, D. H. L. (1982). Molecular basis of bunyavirus per os infection of mosquitoes: role of the middle-sized RNA segment. *Proceedings of National Academy of Sciences* **79**, 1295-1297.

Beaty, B. J., Rayms-Keller, A., Borucki, M. K. & Blair, C. D. (2000). La Crosse encephalitis virus and mosquitoes: a remarkable relationship. *ASM News* **66**, 349-357.

Beaty, B. J., Sundin, D. R., Chandler, L. J. & Bishop, D. H. L. (1985). Evolution of Bunyaviruses by genome reassortment in dually infected mosquitoes (*Aedes triseriatus*). *Science* **230**, 548-550.

Beaty, B. J. & Thompson, W. J. (1975). Emergence of La Crosse virus from endemic foci. *American Journal of Tropical Medicine* **24**, 685-690.

Beck, E. T. (2007). Dissertation. Colorado State University, Fort Collins, CO.

Beck, E. T., Bosio, C. F., Geske, D. A., Blair, C. D., Beaty, B. J. & Black, W. C. (2005). An analysis of gene flow among Midwestern populations of the mosquito *Ochlerotatus triseriatus*. *American Journal of Tropical Medicine and Hygiene* **73**, 534-540.

Bennett, R. S., Ton, D. R., Hanson, C. T., Murphy, B. R. & Whitehead, S. S. (2007). Genome sequence analysis of La Crosse virus and *in vitro* and *in vivo* phenotypes. *Virology Journal* **4**, 41-51.

Berry, W. J., Rowley, W. A. & Christensen, B. M. (1987). Influence of developing *Dirofilaria immitis* on the spontaneous flight activity of *Aedes aegypti*. *Journal of Medical Entomology* **24**, 699-701.

Bird, B. H., Khristova, M. L., Rollin, P. E., Ksiazek, T. G. & Nichol, S. T. (2007). Complete genome analysis of 33 ecologically and biologically diverse Rift Valley Fever virus strains reveals widespread virus movement and low genetic diversity due to recent common ancestry. *Journal of Virology* **81**, 2805-2816.

Bishop, D. H. L. (1979). Genetic potential of bunyaviruses. *Current Topics in Microbiology and Immunology* **86**, 1-33.

Bishop, D. H. L. & Beaty, B. J. (1988). Molecular and biochemical studies of the evolution, infection and transmission of insect bunyaviruses. *Philosophical Transactions of the Royal Society of London* **321**, 463-483.

Bishop, D. H. L., Beaty, B. J. & Shope, R. E. (1980). Recombination and gene coding assignments of bunyaviruses and arenaviruses. *Annals of the New York Academy of Sciences* **354**, 84-106.

Black, W. C. & Salman, M. D. (2005). Molecular techniques for epidemiology and the evolution of arthropod-borne pathogens. In *Biology of Disease Vectors*, 2nd edn, pp. 227-255. Edited by W. C. Marquardt. NY: Elsevier.

Blakqori, G. & Weber, F. (2005). Efficient cDNA-based rescue of La Crosse bunyaviruses expressing or lacking the nonstructural protein NSs. *Journal of Virology* **79**, 10420-10428.

Boni, M. F., Gog, J. R., Andreasen, V. & Feldman, M. W. (2006). Epidemic dynamics and antigenic evolution in a single season of influenza A. *Proceedings of the Royal Acadamy of Biological Sciences* **273**, 1307-1316.

Borucki, M. K., Chandler, L. J., Parker, B. M., Blair, C. D. & Beaty, B. J. (1999). Bunyavirus superinfection and segment reassortment in transovarially infected mosquitoes. *Journal of General Virology* **80**, 3173-3179.

Borucki, M. K., Kempf, B. J., Blair, C. D. & Beaty, B. J. (2001). The Effect of mosquito passage on the La Crosse virus genotype. *Journal of General Virology* **82**, 2919-2926.

Borucki, M. K., Kempf, B. J., Blitvich, B. J., Blair, C. D. & Beaty, B. J. (2002). La Crosse virus: replication in vertebrate and invertebrate hosts. *Microbes and Infection* **4**, 341-350.

Bridgen, A. F., Weber, F., Fazakerley, J. K. & Elliot, R. M. (2001). Bunyamwera bunyavirus nonstructural protein, NSs, in a nonessential gene product that contributes to viral pathogenesis. *Proceedings of National Academy of Sciences* **98**, 664-669.

Briese, T., Bird, B., Kapoor, V., Nichol, S. T. & Lipkin, W. I. (2006). Batai and Ngari viruses: M segment reassortment and association with severe febrile disease outbreaks in East Africa. *Journal of Virology* **80**, 5627-5630.

Brooks, D. R. & Leman, D. A. (1993). Parascript: Parasites and the Language of Evolution. Washington, DC: Smithsonian Institution Press.

Cabradilla, C. D., Holloway, B. P. & Obijeski, J. F. (1983). Molecular cloning and sequencing of the La Crosse virus-S RNA. *Virology* **128**, 463-468.

Calisher, C. H. (1994). Medically important arboviruses of the United States and Canada. *Clinical Microbiology Reviews* **7**, 89-116.

Calisher, C. H. (2005). Persistent emergence of Dengue. *Emerging Infectious Diseases* **11**, 738-739.

CDC (2005). Historical summaries of notifiable diseases in the United States. *Morbidity and Mortality Weekly Report* **52**, 69-85.

Chamberlain, J., Cook, N., Lloyd, G., Mioulet, V., Tolley, H. & Hewson, R. (2005). Co-evolutionary patterns of variation in small and large RNA segments of Crimean-Congo hemorrhagic fever virus. *Journal of General Virology* **86**, 3337-3341.

Chamberlain, R. W. (1980). History of St. Louis encephalitis. In *St. Louis encephalitis*, pp. 3-61. Edited by T. P. Monath. Washington D.C.: American Public Health Association.

Chamberlain, R. W. & Sudia, W. D. (1961). Mechanism of transmission of viruses by mosquitoes. *Annual Review of Entomology* **6**, 371-390.

Chandler, L. J., Beaty, B. J., Baldridge, G. D., Bishop, D. H. L. & Hewlett, M. J. (1990). Heterologous reassortment of bunyaviruses in *Aedes triseriatus* mosquitoes and transovarial and oral transmission of newly evolved genotypes. *Journal of General Virology* **71 (pt 5)**, 1045-1040.

Chandler, L. J., Blair, C. D. & Beaty, B. J. (1998). La Crosse infection of *Aedes triseriatus* ovaries before dissemination of virus from the midgut. *Journal of Medical Entomology* **35**, 567-572.

Chandler, L. J., Hogge, G., Endres, M., Jacoby, D. R., Nathanson, N. & Beaty, B. J. (1991). Reassortment of La Crosse and Tahyna bunyaviruses in *Aedes triseriatus* mosquitoes. *Virus Research* **20**, 181-191.

Chun, R. W. M., Thompson, W. H., Grabow, J. D. & Matthews, C. G. (1968). California arbovirus encephalitis in children. *Neurology* **18**, 369-375.

Clements, A. N. (1992). The Biology of Mosquitoes: Development, Nutrition, and Reproduction, pp. 509. London: Chapman and Hall.

Clerx-van-Haaster, C. M., Akashi, H., Auperin, D. D. & Bishop, D. H. L. (1982). Nucleotide sequence analysis and predicted coding of bunyavirus genome RNA species. *Journal of Virology* **41**, 119-128.

Cockerham, C. C. (1969). Variance of gene frequencies. *Evolution* **23**, 72-83.

Cockerham, C. C. (1973). Analysis of gene frequencies. *Genetics* **74**, 679-700.

Danielova, V. (1974). The property changes of two Tahyna virus strains by influence of serial passages in various environments. *Zentralbl Bakteriol* **229**, 323-333.

de la Torre, J. C., Martinez-Salas, E., Diez, J., Villaverde, A., Gebauer, F., Rocha, E., Davila, M. & Domingo, E. (1988). Coevolution of cells and viruses in a persistent infection of foot and mouth disease virus in cell culture. *Journal of Virology* **62**, 2050-2058.

Defoliart, G. R. (1983). *Aedes triseriatus*: Vector biology in relationship to the persistence of La Crosse virus in endemic foci. In *California Serogroup VIruses*, pp. 89-104. Edited by C. H. Calisher & W. H. Thompson. New York: Alan R. Liss, Inc.

Devenport, M. & Jacobs-Lorena, M. (2005). The peritrophic matrix of hematophagous insects. In *Biology of Disease Vectors*, 2nd edn, pp. 297-310. Edited by W. C. Marquardt. New York: Elsevier.

Deyde, V. M., Khristova, M. L., Rollin, P. E., Ksiazek, T. G. & Nichol, S. T. (2006). Crimean-Congo hemorrhagic fever virus genomics and global diversity. *Journal of Virology* **80**, 8834-8842.

Diaz, F. J., Black, W. C., Farfan-Ale, J. A., Lorono-Pino, M. A., Olson, K. E. & Beaty, B. J. (2006). Dengue virus circulation and evolution in Mexico: A phylogenetic perspective. *Archives of Medical Research* **37**, 760-773.

Diaz, F. J., Farfan-Ale, J. A., Olson, K. E., Lorono-Pino, M. A., Gubler, D. J., Blair, C. D., Black, W. C. & Beaty, B. J. (2002). Genetic variation within the premembrane coding region of dengue viruses from the Yucatan peninsula of Mexico. *American Journal of Tropical Medicine and Hygiene* **67**, 93-101.

Eckels, K. H., Brandt, W. E., Harrison, V. R., McCown, J. M. & Russell, P. K. (1976). Isolation of a temperature-sensitive Dengue-2 virus under conditions suitable for vaccine development. *Infection and Immunity* **14**, 1221-1227.

Eigen, M. (1996). On the nature of viral quasispecies. *Trends in Microbiology* **4**, 212-214.

El Said, L. H., Vorndham, V., Gentsch, J. R., Clewley, J. P., Calisher, C. H., Klimas, R. A., Thompson, W. H., Grayson, M., Trent, D. W. & Bishop, D. H. L. (1979). A comparison of La Crosse isolates obtained from difference ecological niches and an analysis of the structural components of California encephalitis serogroup viruses and other Bunyaviruses. *American Journal of Epidemiology* **28**, 364-386.

Elliott, R. M. (1985). Identification of nonstructural proteins encoded by viruses of the Bunyamwera serogroup (Family: *Bunyaviridae*). *Virology* **143**, 119-126.

Endres, M. J., Jacoby, D. R., Janssen, R. S., Gonzalez-Scarano, F. & Nathanson, N. (1989). The large viral RNA segment of California serogroup bunyaviruses encodes the large viral protein. *Journal of General Virology* **70 (Pt 1)**, 223-228.

Erwin, P. C., Jones, T. F., Gerhardt, R. R., Halford, S. K., Smith, A. B., Patterson, L. E. R., Gottfried, K. L., Burkhalter, K. L., Nasci, R. S. & Schaffner, W. (2002). La Crosse encephalitis in Eastern TN: Clinical, environmental, and entomological characteristics from a blinded cohort study. *American Journal of Epidemiology* **155**, 1060-1065.

Excoffier, L., Smouse, P. & Quattro, J. (1992). Analysis of molecular variance inferred from metric distances among DNA haplotype: Application to human mitochondrial DNA restriction data. *Genetics* **131**, 479-491.

Ferveur, J. F., Cobb, M., Boukella, H. & Jallon, J. M. (1996). Worldwide variation in *Drosophila* sex pheromone: Behavioral effects, genetic bases, and potential evolutionary consequences. *Genetica* **97**, 73-80.

Ferveur, J. F., Savarit, F., O'Kane, C. J., Sureau, G., Greenspan, R. J. & Jallon, J. M. (1997). Genetic feminization of pheromones and its behavioral consequences in *Drosophila* males. *Science* **276**, 1555-1558.

Feuer, R., Boone, J. D., Netski, D., Morzunov, S. P. & St.Jeor, S. C. (1999). Temporal and spatial analysis of Sin Nombre virus quasispecies in naturally infected rodents. *Journal of Virology* **73**, 9544-9554.

Fine, P. E. M. (1975). Vectors and vertical transmission: an epidemiologic perspective. *Annals New York Academy of Sciences* **266**, 173-194.

Fine, P. E. M. & LeDuc, J. W. (1978). Towards a quantitative understanding of the epidemiology of Keystone virus in the Eastern United States. *American Journal of Tropical Medicine and Hygiene* **27**, 322-338.

Fleuriet, A. (1988). Maintenance of a hereditary virus - the Sigma virus in populations of its host, *Drosophila melanogaster*. *Evolutionary Biology* **23**, 1-30.

Fujii, K., Fujii, Y., Noda, T., Muramoto, Y., Watanabe, T., Takada, A., Goto, H., Horimoto, T. & Kawaoka, Y. (2005). Importance of both the coding and the segment-specific noncoding regions of the influenza a virus NS segment for its efficient incorporation into virions. *Journal of Virology* **79**, 3766-3774.

Fujii, Y., Goto, H., Watanabe, T., Yoshida, T. & Kawaoka, T. (2003). Selective incorporation of influenza virus RNA segments into virions. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 2002-2007.

Fuller, F. & Bishop, D. H. L. (1982). Identification of virus coded nonstructural polypeptides in bunyavirus-infected cells. *Journal of Virology* **41**, 643-648.

Gabitzsch, E. S., Blair, C. D. & Beaty, B. J. (2006). Effect of La Crosse virus infection on insemination rates in female *Aedes triseriatus* (Diptera: Culicidae). *Journal of Medical Entomology* **43**, 850-852.

Gauld, L. W., Hanson, R. P., Thompson, W. H. & Sinha, S. K. (1974). Observations on a natural cycle of La Crosse virus (California group) in Southwestern Wisconsin. *American Journal of Tropical Medicine and Hygiene* **23**, 983-992.

Gentsch, J., Wynne, L. R., Clewley, J. P., Shope, R. E. & Bishop, D. H. L. (1977). Formation of recombinants between snowshoe hare and La Crosse bunyaviruses. *Journal of Virology* **24**, 893-902.

Gentsch, J. R. & Bishop, D. H. L. (1979). M viral RNA segment of bunyaviruses codes for 2 glycoproteins, G1 and G2. *Journal of Virology* **30**, 767-770.

Gentsch, J. R., Rozhon, E. J., Klimas, R. A., El Said, L. H., Shope, R. E. & Bishop, D. H. L. (1980). Evidence from recombinant bunyavirus studies that the M RNA gene products elicit neutralizing antibodies. *Virology* **102**, 190-204.

Gerhardt, R. R., Gottfried, K. L., Apperson, C. S., Davis, B. S., Erwin, P. C., Smith, A. B., Panella, N. A., Powell, E. E. & Nasci, R. S. (2001). First isolation of La Crosse virus from naturally infected *Aedes albopictus*. *Emerging Infectious Diseases* **7**.

Gerrard, S. R., Li, L., Barrett, A. D. & Nichol, S. T. (2004). Ngari virus is a Bunyamwera virus reassortant that can be associated with large outbreaks of hemorrhagic fever in Africa. *Journal of Virology* **78**, 8922-8926.

Girard, Y. A., Popov, V., Wen, J., Han, V. & Higgs, S. (2005). Ultrastructural study of West Nile virus pathogenesis in *Culex pipiens quinquefasciatus* (Diptera : Culicidae). *Journal of Medical Entomology* **42**, 429-444.

Girard, Y. A., Schneider, B. S., McGee, C. E., Wen, J., Han, V. C., Popov, V., Mason, P. W. & Higgs, S. (2007). Salivary gland morphology and virus transmission during long-term cytopathologic West Nile virus infection in Culex mosquitoes. *American Journal of Tropical Medicine and Hygiene* **76**, 118-128.

Gonzalez-Scarano, F. (1985). La Crosse virus G1 glycoprotein undergoes a conformational change at the pH of fusion. *Virology* **140**, 209-216.

Gonzalez-Scarano, F., Beaty, B. J., Sundin, D., Janssen, R., Endres, M. J. & Nathanson, N. (1988). Genetic determinants of the virulence and infectivity of La Crosse virus. *Microbial Pathogenesis* **4**, 1-7.

Gonzalez-Scarano, F., Jacoby, D., Griot, C. & Nathanson, N. (1992). Genetics, infectivity, and virulence of California serogroup viruses. *Virus Research* **24**, 123-135.

Gonzalez-Scarano, F., Pobjecky, N. & Nathanson, N. (1987). Role of the G1 protein of La Crosse virus in fusion. In *Biology of negative strand viruses*, pp. 33-39. Edited by M. B. a. K. D.: Elsevier.

Gonzalez-Scarano, F., Shope, R. B., Calisher, C. E. & Nathanson, N. (1982). Characterization of monoclonal antibodies against the G1 and N proteins of La Crosse and Tahyna, two California serogroup bunyaviruses. *Virology* **120**, 42-53.

Grady, L., Sanders, M. L. & Campbell, W. P. (1987). The sequence of the M RNA of an isolate of La Crosse virus. *Journal of General Virology* **68 (part 2)**, 3057-3071.

Graham, D. H., Holmes, J. L., Higgs, S., Beaty, B. J. & Black, W. C. (1999). Selection of refractory and permissive strains of *Aedes triseriatus* (Diptera : Culicidae) for transovarial transmission of La Crosse virus. *Journal of Medical Entomology* **36**, 671-678.

Gratz, N. G. (2004). Critical review of the vector status of *Aedes albopictus*. *Medical and Veterinary Entomology* **18**, 215-227.

Grayson, M. & Calisher, C. H. (1983). California serogroup viruses in New York State: a retrospecitve analysis of subtype distribution patterns and their epidemiologic significance, 1965-1981. In *California Serogroup Viruses*, pp. 257-267. Edited by C. H. Calisher. New York: Liss.

Greenspan, R. J. & Ferveur, J. F. (2000). Courtship in *Drosophila*. *Annual Review of Genetics* **34**, 205-232.

Grimstad, P. R., Barrett, R. L. & Humphrey, R. L. (1984). Serologic evidence for widespread infection with La Crosse and St. Louis encephalitis viruses in the Indiana population. *American Journal of Epidemiology* **119**, 913-930.

Grimstad, P. R., Ross, Q. E. & Craig, G. B. (1980). *Aedes triseriatus* and La Crosse virus II: Modification of mosquito feeding behavior by virus infection. *Journal of Medical Entomology* **17**, 1-7.

Gubler, D. J. (1998). Dengue and Dengue Hemorrhagic Fever. *Clinical Microbiology Reviews* **11**, 480-496.

Gubler, D. J. (2002). The global emergence/resurgence of arboviral diseases as public health problems. *Archives of Medical Research* **33**, 330-342.

Gubler, D. J. (2004). The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? *Comparative Immunology, Microbiology, and Infectious Diseases* **27**, 319-330.

Gunderson, C. B. & Brown, K. L. (1983). Clinical aspects of La Crosse encephalitis, a preliminary report. In *California Serogroup Viruses*, pp. 169-177. Edited by C. Calisher, Thompson, WH. New York: AR Liss.

Hammond, S. M. (2005). Dicing and Slicing: The core machinery of the RNA interference pathway. *Federation of European Biochemical Societies.*

Hardy, J. L. (1988). Susceptibility and resistance of vector mosquitoes. In *The Arboviruses. Epidemiology and Ecology*, pp. 87-126. Edited by T. P. Monath. Boca Raton, FL: CRC Press.

Hartl, D. L. & Clark, A. G. (1997). Principles of Popoulation Genetics, 3rd edn. Sunderland (MA): Sinauer Associates.

Hewson, R., Gmyl, A., Gmyl, L., Smirnova, S. E., Karganova, G., Jamil, B., Hasan, R., Chamberlain, J. & Clegg, C. (2004). Evidence of segment reassortment in Crimean-Congo hemorrhagic fever virus. *Journal of General Virology* **85**, 3059-3070.

Higgs, S. (2004). How do mosquitoes live with their viruses? In *Microbe-Vector Interactions in Vector-Borne Diseases*, pp. 103-137. Edited by S. H. Gillespie, G. L. Smith & A. Osbourn: Cambridge University Press.

Higgs, S. & Beaty, B. J. (2005). Natural Cycles of Vector-borne Pathogens. In *Biology of Disease Vectors*, 2nd edn, pp. 167-185. Edited by W. Marquardt. New York: Elsevier.

Hill, W. G. & Robertson, A. (1966). Effect of linkage on limits to artificial selection. *Genetic Research* **8**, 269-294.

Holland, J., Spindler, K., Horodyski, F., Grabu, E., Nichol, S. T. & VandePol, S. (1982). Rapid evolution of RNA viruses. *Science* **215**, 1577-1585.

Huang, C., Shope, R. E., Spargo, B. & Campbell, W. P. (1996). The S RNA genomic sequences of Inkoo, San Angelo, Serra de Navio, South River and Tahyna bunyaviruses. *Journal of General Virology* **77**, 1761-1768.

Huang, C., Thompson, W. H. & Campbell, W. P. (1995). Comparison of the M RNA genome segments of 2 human isolates of La Crosse virus. *Virus Research* **36**, 177-185.

Huang, C., Thompson, W. H., Karabatos, N., Grady, L. & Campbell, W. P. (1997). Evidence that fatal human infections with La Crosse virus may be associated with a narrow range of genotypes. *Virus Research* **48**, 143-8.

Hubalek, Z. & Halouzka, J. (1999). West Nile fever: a reemerging mosquito-borne viral disease in Europe. *Emerging Infectious Diseases* **5**, 643-650.

Jallon, J. M. (1984). A few chemical words exchanged by *Drosophila* during courtship and mating. *Behavioral Genetics* **14**, 112-121.

Janssen, R., Gonzalez-Scarano, F. & Nathanson, N. (1984). Mechanisms of bunyavirus virulence - Comparative pathogenesis of a virulent strain of La Crosse and an avirulent strain of Tahyna virus. *Laboratory Investigation* **50**, 447-455.

Janssen, R., Nathanson, N., Endres, M. & Gonzalez-Scarano, F. (1986). Virulence of La Crosse virus is under polygenic control. *Journal of Virology* **59**, 1-7.

Johnson, R. (1983). Pathogenesis of La Crosse virus in mice. *Progressive Clinical Biological Research* **123**, 139-144.

Jones, T. F., Craig, A. S., Nasci, R. S., Patterson, L. E. R., Erwin, P. C., Gerhardt, R. R., X.T., U. & Schaffner, W. (1999). Newly recognized focus of La Crosse encephalitis in Tennessee. *Clinical Infectious Diseases* **28**, 93-97.

Kalfayan, B. (1983). Pathology of La Crosse virus infection in humans. In *California Serogroup Viruses*, pp. 179-186. Edited by C. H. a. T. Calisher, WH. New York: Alan R. Liss, Inc.

Kappus, K. D., Calisher, C. H. & Matthews, J. H. (1982). La Crosse virus infection and disease in western North Carolina. *American Journal of Tropical Medicine and Hygiene* **31**, 556-560.

Kappus, K. D., Monath, T. P., Kaminski, R. M. & Calisher, C. H. (1983). Reported encephalitis associated with California serogroup virus infection in the United States. In *California Serogroup Viruses*, pp. 31-41. Edited by C. H. Calisher & W. H. Thompson. New York: A.R. Liss.

Karabatsos, N. (1985. 2001 update). International Catalogue of Arboviruses, Including Certain Other Viruses of Vertebrates. *American Society of Tropical Medicine and Hygiene*.

Karber, G. (1931). Beitrag zur kollektiven Behandlung pharmakologischer Reihenversuche. *Arch Exp Pathol Parmakol* **162**, 480-483.

Kempf, B. J., Blair, C. D. & Beaty, B. J. (2006). Quantitative analysis of La Crosse virus transcription and replication in cell cultures and mosquitoes. *American Journal of Tropical Medicine and Hygiene* **74**, 224-232.

Kimura, M. (1969). The rate of molecular evolution considered from the standpoint of population genetics. *Proceedings of National Academy of Sciences* **63**, 1181-1188.

Klimas, R. A., Thompson, W. H., Calisher, C. H., Clark, G. G., Grimstad, P. R. & Bishop, D. H. L. (1981). Genotypic varieties of La Crosse virus isolated from different geographic regions of the continental United States and evidence for a naturally occurring intertypic recombinant La Crosse virus. *American Journal of Epidemiology* **114**, 112-131.

Klowden, M. J. & Zwiebel, L. J. (2005). Vector Olfaction and Behavior. In *Biology of Disease Vectors*, 2nd edn, pp. 277-288. Edited by W. C. Marquardt. NY: Elsevier.

Kohl, A., Clayton, R. F., Weber, F., Bridgen, A. F., Randall, R. E. & Elliot, R. M. (2003). Bunyamwera virus nonstructural NSs counteracts interferon regulatory factor 3-mediated induction of early cell death. *Journal of Virology* **77**, 7999-8008.

Kolakofsky, D., and D. Hacker (1991). Bunyavirus RNA synthesis: genome transcription and replication. *Current Topics in Microbiology and Immunology* **169**, 143-159.

Lanciotti, R. S., Roehrig, J. T., Deubel, V., Smith, J., Parker, M., Steele, K., Crise, B., Volpe, K. E., Crabtree, M. B., Scherret, J. H., Hall, R. A., MacKenzie, J. S., Cropp, C. B., Panigraphy, B., Ostlund, E., Schmitt, B., Malkinson, M., Banet, C., Weissman, J., Komar, N., Savage, H. M., Stone, W., McNamara, T. & Gubler, D. J. (1999). Origin of the West Nile virus responsible for an outbreak of encephalitis in the northeastern United States. *Science* **286**, 2333-2337.

Laver, W. G., Webster, R. G. & Chu, C. M. (1984). From the National Institutes of Health. Summary of a meeting on the origin of pandemic influenza viruses. *Journal of Infectious Diseases* **149**, 108-115.

LeDuc, J. W. (1979). The ecology of California group viruses. *Journal of Medical Entomology* **16**, 1-7.

Lee, J. H., Rowley, W. A. & Platt, K. B. (2000). Longevity and spontaneous flight activity of *Culex tarsalis* infected with Western equine encephalomyelitis virus. *Journal of Medical Entomology* **37**, 187-193.

Leopold, R. A. & Degrugillier, M. E. (1973). Sperm penetration of housefly eggs: evidence for involvement of a female accessory secretion. *Science* **181**, 555-556.

Li, D., Schmaljohn, A. L., Anderson, K. & Schmaljohn, C. S. (1995a). Complete nucleotide sequences of the M and S segments of two hantavirus isolates from California: Evidence for reassortment in nature among viruses related to Hantavirus Pulmonary Syndrome. *Virology* **206**, 973-983.

Li, D. X., Schmaljohn, A. L., Anderson, K. & Schmaljohn, C. S. (1995b). Complete nucleotide sequences of the M segment and S segment of 2 Hantavirus isolates from California - Evidence for reassortment in nature among viruses related to Hantavirus pulmonary syndrome. *Virology* **206**, 973-983.

Lisitza, M. A., DeFoliart, G. R., Yuill, T. M. & Karandinos, M. G. (1977). Prevalence rates of La Crosse virus (California encephalitis group) in larvae from overwintered eggs of *Aedes triseriatus*. *Mosquito News* **37**, 745-750.

Long, J. C. (1986). The allele correlation structure of Gainj and Kalam speaking people. I. The estimation and interpretation of Wright's F-statistics. *Genetics* **112**, 629-647.

Lorono-Pino, M. A., Farfan-Ale, J. A., Zapata-Peraza, A. L., Rosado-Paredes, E. P., Flores-Flores, L. F., Garcia-Rejon, J. E., Diaz, F. J., Blitvich, B. J., Andrade-Narvaez, M., Jimenez-Rios, E., Blair, C. D., Olson, K. E., Black, W. & Beaty, B. J. (2004). Introduction of the American/Asian genotype of dengue 2 virus into the Yucatan State of Mexico. *American Journal of Tropical Medicine and Hygiene* **71**, 485-492.

Ludwig, G. V., Christensen, B. M., Yuill, T. M. & Schultz, K. T. (1989). Enzyme processing of La Crosse virus glycoprotein G1 - a bunyavirus-vector infection model. *Virology* **171**, 108-113.

Ludwig, G. V., Israel, B. A., Christensen, B. M., Yuill, T. M. & Schultz, K. T. (1991). Role of La Crosse virus glycoproteins in attachment of virus to host cells. *Virology* **181**, 564-571.

MacDonald, G. (1957). Epidemiology and Control of Malaria. London: Oxford University Press.

Marfin, A. A. & Gubler, D. J. (2001). West Nile encephalitis: an emerging disease in the United States. *Clinical Infectious Diseases* **33**, 1713-1719.

Mather, T. N. & Defoliart, G. R. (1984). Conditions for increased insemination rates in caged *Aedes triseriatus* (Diptera: Culicidae) reared from field-collected eggs. *American Journal of Tropical Medicine and Hygiene* **33**, 731-735.

McGaw, M. M., Chandler, L. J., Wasieloski, L. P., Blair, C. D. & Beaty, B. J. (1998). Effect of La Crosse virus infection on overwintering of *Aedes triseriatus*. *American Journal of Tropical Medicine* **58**, 168-175.

McJunkin, J. E., de los Reyes, E. C., Irazuzta, J. E., Caceres, M. J., Khan, R. R., Minnich, L. L., Fu, K. D., Lovett, G. D., Tsai, T. & Thompson, A. (2001). La Crosse encephalitis in children. *New England Journal of Medicine* **344**, 801-807.

Meyer, K. F., Haring, C. M. & Howitt, B. (1931). The etiology of epizootic encephalomyelitis of horses in the San Joaquin Valley. *Science* **74**, 227-228.

Miller, B. R., Beaty, B. J. & Lorenz, L. H. (1982). Variation of La Crosse virus filial infection rates in geographic strains of *Aedes triseriatus* (Diptera: Culicidae). *Journal of Medical Entomology* **19**, 213-214.

Miller, B. R., DeFoliart, G. R. & Yuill, T. M. (1977). Vertical transmission of La Crosse virus (California encephalitis group): Transovarial and filial infection rates in Aedes triseriatus. *Journal of Medical Entomology* **14**, 437-440.

Miller, B. R., DeFoliart, G. R. & Yuill, T. M. (1979). *Aedes triseriatus* and La Crosse virus: lack of infection in eggs of the first ovarian cycle following oral infection of females. *American Journal of Tropical Medicine* **28**, 897-901.

Mitchell, C. J. (1983). Mosquito vector competence and arboviruses. In *Current Topics in Vector Research*, pp. 63-92. Edited by K. F. Harris. N.Y.: Praeger Publishers.

Monath, T. P. (1990). Flaviviruses. In *Virology*, pp. 763-814. Edited by D. K. B.N. Fields, et al. New York: Raven Press.

Monath, T. P., Nuckolls, J. G., Berall, J., Bauer, H., Chappell, W. A. & Coleman, P. H. (1970). Studies on California encephalitis in Minnesota. *American Journal of Epidemiology* **92**, 40-50.

Mumford, J. A. (2007). Vaccines and viral antigenic diversity. *Revue Scientifique et Technique* **26**, 69-90.

Muramoto, Y., Takada, A., Fujii, K., Noda, T., Iwatsuki-Horimoto, K., Watanabe, S., Horimoto, T., Kida, H. & Kawaoka, Y. (2006). Hierarchy among viral RNA (vRNA) segments in their role in vRNA incorporation into influenza A virions. *Journal of Virology* **80**, 2318-2325.

Novella, I. S., Clarke, D. K., Quer, J., Duarte, E. A., Lee, C. H., Weaver, S. C., Elena, S. F., Moya, A., Domingo, E. & Holland, J. (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.

Nunes, M. R. T., Travassos de Rosa, A. P. A., Weaver, S. C., Tesh, R. B. & Vasconcelos, P. F. C. (2005). Molecular epidemiology of Group C viruses isolated in the Americas. *Journal of Virology* **79**, 10561-10570.

Obijeski, J. F., McCauley, J. & Skehel, J. J. (1980). Nucleotide sequences at the terminal of La Crosse virus RNAs. *Nucleic Acids Research* **8**, 2431-2438.

Overby, A. K., Pettersson, R. F. & Neve, E. P. A. (2007). The glycoprotein cytoplasmic tail of Uukuniemi virus (Bunyaviridae) interacts with ribonucleoproteins and is critical for genome packaging. *Journal of Virology* **81**, 3198-3205.

Palese, P. & Young, J. F. (1982). Variation of influenza A, B, and C viruses. *Science* **215**, 1468-1474.

Pantuwatana, S., Thompson, W. H., Watts, D. M. & Hanson, R. P. (1972). Experimental infection of chipmunks and squirrels with La Crosse and Trivitattus viruses and biological transmission of La Crosse virus by *Aedes triseriatus*. *American Journal of Tropical Medicine* **21**, 476-481.

Pantuwatana, S., Thompson, W. H., Watts, D. M., Yuill, T. M. & Hanson, R. P. (1974). Isolation of La Crosse virus from field collected *Aedes triseriatus* larvae. *American Journal of Tropical Medicine* **23**, 246-250.

Parsonson, I. M. & McPhee, D. A. (1985). Bunyavirus pathogenesis. *Advanced Virus Research* **30**, 279-316.

Patrican, L. A., DeFoliart, G. R. & Yuill, T. M. (1985). La Crosse viremias in juvenile, subadult, and adult chipmunks following feeding by transovarially infected *Aedes triseriatus*. *American Journal of Tropical Medicine* **34**, 569-602.

Platt, K. B., Linthicum, K. J., Myint, K. S. A., Innis, B. L., Lerdthusnee, K. & Vaughn, D. W. (1997). Impact of dengue virus infection on feeding behavior of *Aedes aegypti*. *American Journal of Tropical Medicine and Hygiene* **57**, 119-125.

Polerstock, A. R., Eigenbrode, S. D. & Klowden, M. J. (2002). Mating alters the cuticular hydrocarbons of female *Anopheles gambiae* sensu stricto and *Aedes aegypti*. *Journal of Medical Entomology* **39**, 545-552.

Pomonis, J. G. (1989). Cuticular hydrocarbon of the screwworm, *Cochlimyia hominivorax*. Isolation, identification, and quantification as a function of age, sex, and irradiation. *Journal of Chemical Ecology* **15**, 2301-2317.

Porter, C. H. & Defoliart, G. R. (1985). Observations on physiological age, insemination, and *Ascogregarina* infection in a southern Wisconsin population of *Aedes triseriatus*. *Journal of American Mosquito Control Association* **1**, 238-240.

Reisen, W. K. (2003). Epidemiology of St. Louis encephalitis virus. *Advances in Virus Research* **61**, 139-183.

Rico-Hesse, R. (1990). Molecular evolution and distribution of dengue viruses type 1 and 2 in nature. *Virology* **173**, 479-493.

Rust, R. S., Thompson, W. H., Matthews, C. G., Beaty, B. J. & Chun, R. W. M. (1999). La Crosse and other forms of California encephalitis. *Journal of Child Neurology* **14**, 1-14.

Rutledge, J. C., Ward, R. A. & Gould, D. J. (1964). Studies on the feeding response of mosquitoes to nutritive solutions in a new membrane feeder. *Mosquito News* **24**, 407-420.

Sanchez-Vargas, I., Travanty, E. A., Keene, K. M., Franz, A. W., Beaty, B. J., Blair, C. D. & Olson, K. E. (2004). RNA interference, arthropod-borne viruses, and mosquitoes. *Virus Research* **102**, 65-74.

Schmaljohn, C. S. (1996). *Bunyaviridae*: the viruses and their replication.

Schmaljohn, C. S. & Hooper, J. W. (2001). *Bunyaviridae*: The Viruses and Their Replication. In *Field's Virology*, pp. 1581-1602. Edited by D. M. a. H. Knipe, P.M. Philadelphia: Lippincott, Williams and Wilkins.

Scholl, P. J., Porter, C. H. & Defoliart, G. R. (1979). *Aedes triseriatus*: Persistence of nulliparous females under field conditions. *Mosquito News* **39**, 368-371.

Scott, T. W., Hildreth, S. W. & Beaty, B. J. (1984). The distribution and development of Eastern Equine encephalitis virus in its enzootic mosquito vector, *Culiseta melanura*. *American Journal of Tropical Medicine and Hygiene* **33**, 300-310.

Scott, T. W. & Lorenz, L. H. (1998). Reduction of *Culiseta melanura* fitness by eastern equine encephalomyelitis virus. *American Journal of Tropical Medicine and Hygiene* **59**, 341-346.

Scott, T. W. & Weaver, S. C. (1989). Eastern Equine Encephalomyelitis Virus - Epidemiology and Evolution of Mosquito Transmission. *Advances in Virus Research* **37**, 277-328.

Seecof, R. (1966). Low titers in *Drosophila* that regularly transmit the Sigma virus infection. *Genetics* **53**, 79-83.

Seecof, R. (1968). The Sigma virus infection of *Drosophila melanogaster*. *Current Topics in Microbiology and Immunology* **42**, 59-93.

Seecof, R. L. (1964). Deleterious Effects on *Drosophila* Development Associated with Sigma Virus Infection. *Virology* **22**, 142-&.

Sejvar, J. J. (2006). The evolving epidemiology of viral encephalitis. *Current Opinion in Neurology* **19**, 350-357.

Seymour, C., Amundson, T. E., Yuill, T. M. & Bishop, D. H. L. (1983). Experimental infection of chipmunks and snowshoe hares with La Crosse and snowshoe hare viruses and four of their reassortants. *American Journal of Tropical Medicine and Hygiene* **32**, 1147-1153.

Shope, R. E. (1985). Bunyaviruses. In *Virology*, pp. 1055-1083. Edited by B. N. Fields. New York: Raven Press.

Shope, R. E., Rozhon, E. J. & Bishop, D. H. L. (1981). Role of the middle sized bunyavirus segment in mouse virulence. *Virology* **114**, 273-276.

Shroyer, D. A. & Craig, G. B. (1980). Egg hatchability and diapause in *Aedes triseriatus* (Diptera, Culicidae) - Temperature induced and photoperiod-induced latencies. *Annals of the Entomological Society of America* **73**, 39-43.

Snodgrass, R. E. (1959). The anatomical life of the mosquito. In *Smithsonian Miscellaneous Collections*. Washington: Smithsonian Institution.

Soldan, S. S. & Gonzalez-Scarano, F. (2005). Emerging infectious diseases: The Bunyaviridae. *Journal of Neurovirology* **11**, 412-423.

Soldan, S. S., Plassmeyer, M. L., Matukonis, M. K. & Gonzalez-Scarano, F. (2005). La Crosse virus nonstructural protein NSs counteracts the effects of short interfering RNA. *Journal of Virology* **79**, 234-244.

Streitenfeld, H., Boyd, A., Fazakerley, J. K., Bridgen, A. F., Elliot, R. M. & Weber, F. (2003). Activation of PKR by Bunyamwera virus is independent of the viral interferon antagonist NSs. *Journal of Virology* **77**, 5507-5511.

Sundin, D. R., Beaty, B. J., Nathanson, N. & Gonzalez-Scarano, F. (1987). A G1 glycoprotein epitope of La Crosse virus - a determinant of infection of *Aedes triseriatus*. *Science* **235**, 591-593.

Swofford, D. (2003). PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods), 4 edn. Sutherland, MA: Sinauer Associates.

Tatem, A. J., Hay, S. I. & Rogers, D. J. (2006). Global traffic and disease vector dispersal. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 6242-6247.

Ten Broeck, C. & Merrill, M. H. (1933). A serological difference between Eastern and Western equine encephalomyelitis. *Proceedings of the Society for Experimental Biology and Medicine* **31**, 217-220.

Tesh, R. (1980). Vertical transmission of arthropod-borne viruses of vertebrates. In *Vectors of Disease Agents*. Edited by J. J. McKelvey & R. F. Harwood. New York: Praeger Scientific Publishers.

Tesh, R. & Shroyer, D. A. (1980). The mechanism of arbovirus transovarial transmission in mosquitoes: San Angelo virus in *Aedes albopictus*. *American Journal of Tropical Medicine and Hygiene* **29**, 1394-1404.

Thompson, W. H. (1979). Higher venereal infection and transmission rates with La Crosse virus in *Aedes triseriatus* engorged before mating. *American Journal of Tropical Medicine and Hygiene* **28**, 890-896.

Thompson, W. H. & Beaty, B. J. (1977). Venereal transmission of La Crosse (California encephalitis) arbovirus in *Aedes triseriatus* mosquitoes. *Science* **196**, 530-531.

Thompson, W. H. & Beaty, B. J. (1978). Venereal transmission of La Crosse virus from male to female *Aedes triseriatus*. *American Journal of Tropical Medicine and Hygiene* **27**, 187-196.

Thompson, W. H., Kalfayan, B. & Anslow, R. O. (1965). Isolation of California Encephalitis group virus from a fatal human illness. *American Journal of Epidemiology* **81**, 245-253.

Trabalon, M., Campan, M., Clement, J. L., Thon, B., Lange, C. & Leferve, J. (1988). Changes in cuticular hydrocarbon composition in relation to age and sexual behavior in the female *Calliphora vomitoria*. *Behavioral Processes* **17**, 107-115.

Trent, D. W., Grant, J. A., Vorndam, A. V. & Monath, T. P. (1981). Genetic heterogeneity among Saint Louis encephalitis v irus isolates of different geographic origin. *Virology* **114**, 319-332.

Tsai, T. F. (1991). Arboviral infections in the United States. *Infectious Disease Clinic of North America* **5**, 73-102.

Turell, M. J., Hardy, J. L. & Reeves, W. C. (1982). Stabilized infection of California encephalitis virus in *Aedes dorsalis*, and its implications for viral maintenance in nature. *American Journal of Tropical Medicine and Hygiene* **31**, 1252-1259.

Urquidi, V. & Bishop, D. H. L. (1992). Non-random reassortment between the tripartite RNA genomes of La Crosse and snowshoe hare viruses. *Journal of General Virology* **73**, 2255-2265.

Vapalahti, O., Plyusnin, A., Cheng, Y., Manni, T., Brummer-Korvenkontio, M. & Vaheri, A. (1996). Inkoo and Tahyna, the European California serogroup bunyaviruses: sequence and phylogeny of the S RNA segment. *Journal of General Virology* **77**, 1769-1774.

Watanabe, T., Watanabe, S., Noda, T., Fujii, Y. & Kawaoka, Y. (2003). Exploitation of nucleic acid packaging signals to generate a novel influenza virus-based vector stably expressing two foreign genes. *Journal of Virology* **77**, 10575-10583.

Watts, D. M., Morris, C. D., Wright, R. E., DeFoliart, G. R. & Hanson, R. P. (1972). Transmission of La Crosse virus by the mosquitoes *Aedes triseriatus*. *Journal of Medical Entomology* **9**, 125-127.

Watts, D. M., Pantuwatana, S., DeFoliart, G. R., Yuill, T. M. & Thompson, W. H. (1973). Transovarial transmission of La Crosse virus in the mosquito, *Aedes triseriatus*. *Science* **182**, 1140-1141.

Watts, D. M., Thompson, W. H., Yuill, T. M., DeFoliart, G. R. & Hanson, R. P. (1974). Overwintering of La Crosse virus in *Aedes triseriatus*. *American Journal of Tropical Medicine* **23**, 694-700.

Weaver, S. C., Scott, T. W., Lorenz, L. H., Lerdthusnee, K. & Romoser, W. S. (1988). Togavirus associated pathologic changes in the midgut of a natural mosquito vector. *Journal of Virology* **62**, 2083-2090.

Weber, F., Bridgen, A. F., Fazakerley, J. K., Streitenfeld, H., Kessler, N., Randall, R. E. & Elliot, R. M. (2002). Bunyamwera bunyavirus nonstructural protein NSs counteracts the induction of alpha/beta interferon. *Journal of Virology* **76**, 7949-7955.

Weber, F. & Elliot, R. M. (2002). Antigenic drift, antigenic shift and interferon antagonists: how bunyaviruses counteract the immune system. *Virus Research* **88**, 129-136.

Weir, B. S. & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

WHO (1999). Prevention and control of dengue and dengue hemorrhagic fever. Comprehensive guidelines. In *SEARO*, pp. 134. New Delhi, India: WHO Regional Publication.

WHO (2002). Fact Sheet: dengue and dengue hemorrhagic fever. World Health Organization, Geneva.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics* **1**, 80-83.

Woodring, J., Chandler, L. J., Oray, C. T., McGaw, M. M., Blair, C. D. & Beaty, B. J. (1998). Short Report: Diapause, transovarial transmission, and filial infection rates in geographic strains of La Crosse virus infected *Aedes triseriatus*. *American Journal of Tropical Medicine* **58**, 587-588.

Woodruff, B. A., Baron, R. C. & Tsai, T. F. (1992). Symptomatic La Crosse virus infections of the central nervous system: a study of risk factors in an endemic area. *American Journal of Epidemiology* **136**, 320-327.

Yamakawa, M., Yanase, T., Kato, T. & Tsuda, T. (2006). Chronological and geographical variations in the small RNA segment of the teratogenic Akabane virus. *Virus Research* **121**, 84-92.

Yanase, T., Kato, T., Yamakawa, M., Takayoshi, K., Nakamura, K., Kokuba, T. & Tsuda, T. (2006). Genetic characterization of Batai virus indicates a genomic reassortment between orthobunyaviruses in nature. *Archives of Virology* **151**, 2253-2260.