DISSERTATION

GENOMIC CHARACTERIZATION, DETECTION AND MOLECULAR EVOLUTION OF ARTHROPOD-BORNE VIRUSES OF THE FAMILY *BUNYAVIRIDAE*

Submitted by

Amy Julia Lambert

Department of Microbiology, Immunology and Pathology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2010

Doctoral Committee:

Department Head: Sandra Quackenbush

Advisor: Carol Blair

Robert Lanciotti
Chaoping Chen
Barry Miller
William Black IV

ABSTRACT OF DISSERTATION

GENOMIC CHARACTERIZATION, DETECTION AND EVOLUTION OF ARTHROPOD-BORNE VIRUSES OF THE FAMILY *BUNYAVIRIDAE*

The genomic characterization, detection and evolution of arthropod-borne human pathogens and related viruses of the family *Bunyaviridae* are presented. This study began with the determination of primary nucleotide sequence data for a diversity of bunyaviruses of interest that were not previously characterized at the genetic level. Following molecular characterization, an RT-PCR strategy was designed, according to previously determined and newly derived nucleotide sequence data, to target S genomic segments of 47 viruses, including 29 arthropod-borne human pathogens, of the family *Bunyaviridae*. Following development, this strategy was used, in some instances with an expanded capacity for the detection of multiple segments of the bunyavirus genome, for the identification of arthropod-borne bunyaviruses of medical importance and novel circumstance at a global level. Phylogenetic analyses performed on nucleotide sequence data generated by these efforts facilitate an ancestral understanding of the discovered agents. Finally, advanced molecular and phylogenetic analyses of a subset of newly derived sequences are described here to facilitate an enhanced understanding of the evolution of *Bunyaviridae.*

ACKNOWLEDGEMENTS

I would like to thank my graduate committee for supporting me in this effort and for challenging me to grow as a research scientist. I especially thank my mentor and friend, Dr. Robert Lanciotti for giving me this opportunity and for allowing me the flexibility to succeed and fail on my own. Rob, it is my honor and privilege to work for someone of such great professional capability and character; thank you. I'd also like to give special thanks to my main advisor, Dr. Carol Blair for her phenomenal enthusiasm, encouragement, knowledge of virology and investment of time in me. Dr. Blair, you are an incredible role model; thank you. In addition, I thank Dr. Jeff Chang who is a constant inspiration to me as a creative intellect and generous mentor. Dr. Chang, you are truly one-of-a-kind; thank you.

My family has also been endlessly supportive. There is nothing more fulfilling than coming home to my loving husband, Brad and beautiful, vibrant, funny, caring daughter, Julia after a long day. Brad, I've been able to focus on my research because you have taken such wonderful care of Julia while managing your own career as a research scientist; thank you. I also thank my sister, Melinda and her family who have been a second family to Julia while I've worked toward my academic and professional goals. Melinda, Walker, Anna and Frank, thank you for treating my little girl as your own and for giving her so much love. In addition, I thank my brother, Chad for his humor and wisdom. Chad, you've made it easier for me to navigate this process with your refreshing perspective; thank you. Finally, I thank my parents, Fred and Ardith Kerst for making me who I am. Daddy, you are *the* example of hard work, integrity and success in science,

business and life; thank you for showing me how to succeed while maintaining my values. Mommy, you are the most intelligent, generous, caring, compassionate, loving, beautiful person that I have ever known; thank you for being such a wonderful mother and friend.

To my mother

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Classification and overview of arthropod-borne animal viruses of the family *Bunyaviridae*

The family *Bunyaviridae* includes more than 300 distinct members, with over 60 viruses associated with human illness, predominantly organized into four genera: *Hantavirus, Orthobunyavirus, Phlebovirus and Nairovirus,* according to structural, genetic and antigenic characteristics (Barrett and Shope, 2005; Nichol et al, 2005). Exemplifying the great diversity of species classified within the family, the fifth genus of the family *Bunyaviridae*, the genus *Tospovirus,* contains viruses that are known to cause disease in plants, not animals (Nichol et al, 2005). Further illustrating this extraordinary diversity, the taxonomical hierarchy of *Bunyaviridae* is complex, with members of each genus additionally classified into individual groups, subtypes and complexes according to serological characteristics (Barrett and Shope, 2005).

Most viruses of the family *Bunyaviridae* are arthropod-borne with small mammals typically serving as amplifying vertebrate hosts and vectors of transmission including mosquito, tick, sand fly and thrips species. However, members of the genus *Hantavirus* are not associated with an arthropod vector and are thought to be transmitted from infected rodents to humans through aerosolized excreta. This dissertation focuses on analyses of the three arthropod-borne genera of *Bunyaviridae* within which human pathogens are classified: the *Orthobunyavirus, Phlebovirus* and *Nairovirus* genera (Table 1.1). Viruses of these genera are associated with significant human illness worldwide and cause a range of clinical manifestations including severe pediatric encephalitis (associated with La Crosse virus of the genus *Orthobunyavirus*), hemorrhagic fever (associated with Crimean-Congo hemorrhagic fever virus of the genus *Nairovirus*),

Table 1.1 Arthopod-borne human pathogens of the *Orthobunyavirus, Phlebovirus* and *Nariovirus* genera of the family *Bunyaviridae*

Genus	***Serogroup***	***Virus***	***Principal vector***	***Geography***
Orthobunyavirus	*Bunyamwera*	Tacaiuma*	*Culicidae*	S America
Orthobunyavirus	*Bunyamwera*	Bunyamwera	*Culicidae*	Africa
Orthobunyavirus	*Bunyamwera*	Cache Valley*	*Culicidae*	N America
Orthobunyavirus	*Bunyamwera*	Fort Sherman*	*Culicidae*	C America
Orthobunyavirus	*Bunyamwera*	Garissa (Ngari)	?	Africa
Orthobunyavirus	*Bunyamwera*	Germiston*	*Culicidae*	Africa
Orthobunyavirus	*Bunyamwera*	Iiesha*	*Culicidae*	Africa
Orthobunyavirus	*Bunyamwera*	Shokwe*	*Culicidae*	Africa
Orthobunyavirus	*Bunyamwera*	Wyeomia*	*Culicidae*	S America
Orthobunyavirus	*Bunyamwera*	Xingu*	*Culicidae*	S America
Orthobunyavirus	*Bwamba*	Bwamba*	*Culicidae*	Africa
Orthobunyavirus	*Bwamba*	Pongola*	*Culicidae*	Africa
Orthobunyavirus	*Group C*	Apeu*	*Culicidae*	S America
Orthobunyavirus	*Group C*	Caraparu*	*Culicidae*	The Americas
Orthobunyavirus	*Group C*	Itaqui*	*Culicidae*	S America
Orthobunyavirus	*Group C*	Madrid*	*Culicidae*	N America
Orthobunyavirus	*Group C*	Marituba*	*Culicidae*	S America
Orthobunyavirus	*Group C*	Murutucu*	*Culicidae*	S America
Orthobunyavirus	*Group C*	Nepuyo*	*Culicidae*	S+N Am.
Orthobunyavirus	*Group C*	Oriboca*	*Culicidae*	S America
Orthobunyavirus	*Group C*	Ossa*	*Culicidae*	N America
Orthobunyavirus	*Group C*	Restan*	*Culicidae*	S America
Orthobunyavirus	*California*	California Encephalitis	*Culicidae*	N America
Orthobunyavirus	*California*	Guaroa*	*Culicidae*	The Americas
Orthobunyavirus	*California*	Jamestown Canyon*	*Culicidae*	N America
Orthobunyavirus	*California*	Inkoo*	*Culicidae*	N Hemisphere
Orthobunyavirus	*California*	La Crosse	*Culicidae*	N America
Orthobunyavirus	*California*	Snowshoe hare*	*Culicidae*	N America
Orthobunyavirus	*California*	Tahyna*	*Culicidae*	Europe
Orthobunyavirus	*Guama*	Catu*	*Culicidae*	S America
Orthobunyavirus	*Guama*	Guama*	*Culicidae*	The Americas
Orthobunyavirus	*Nyando*	Nyando*	*Culicidae*	Africa
Orthobunyavirus	*Simbu*	Oropouche	*Culicidae, Ceratopogonidae*	S America
Nairovirus	*CCHF*	CCHF	*Ixodidae*	E Hemisphere
Nairovirus	*NSD*	Dugbe	*Ixodidae, Ceratopogonidae*	Africa
Nairovirus	*NSD*	Nairobi sheep disease*	*Ixodidae, Culicidae*	Africa, Asia
Phlebovirus	*Sand Fly*	Alenquer*	*Psychodidae*	S America
Phlebovirus	*SF*	Candiru*	?	S America
Phlebovirus	*SF*	Punta Toro	*Psychodidae*	The Americas
Phlebovirus	*SF*	Rift Valley Fever	*Culicidae*	Africa
Phlebovirus	*SF*	Sandfly fever Naples*	*Psychodidae*	Europe
Phlebovirus	*SF*	Toscana	*Psychodidae*	Europe
Phlebovirus	*SF*	Chagres*	?	C America
Phlebovirus	*SF*	Sandfly fever Sicillian*	*Psychodidae*	Europe

*Indicates either a lack of primary sequence data in the Fall of 2006. List of bunyaviruses adapted from Barrett & Shope, 2005. CCHF, Crimean-Congo hemorrhagic fever; NSD, Nairobi sheep disease.

and retinitis, encephalitis and hemorrhagic fever (associated with Rift Valley fever virus of the genus *Phlebovirus*).

Discovery, history and clinical significance of arthropod-borne human pathogens of the family *Bunyaviridae*

Orthobunyavirus genus

The genus *Orthobunyavirus* is the largest of the family *Bunyaviridae*, with over 160 predominantly mosquito-borne viruses identified as members (Nichol et al, 2005) and a variety of vertebrate species serving as amplifying hosts. The type species, Bunyamwera virus of the Bunyamwera serogroup, was first isolated from *Aedes* mosquitoes during a yellow fever study in Uganda, 1943 (Smithburn et al., 1946). This isolation ultimately led to the discovery of the family *Bunyaviridae*. Subsequent studies occurring in the Americas during the 1950s generated isolates now classified within the genus *Orthobunyavirus* that were antigenically distinct from the classical arbovirus A and B groups (Casals & Whitman, 1961). This distinction resulted in the classification of these viruses as "Group C" viruses (Casals & Whitman, 1961). Additionally, the type member of the California serogroup of the genus *Orthobunyavirus*, California encephalitis virus was isolated in 1941 and subsequently associated with cases of encephalitis in California (Hammon & Reeves, 1952). These early isolations reflect the global distribution of viruses of the genus *Orthobunyavirus*. Their geographic range encompasses a diversity of tropical (e.g., Group C viruses), temperate (e.g., La Crosse and California encephalitis

viruses) and arctic (e.g., Inkoo, Northway viruses) ecologies. Approximately 25 orthobunyaviruses are associated with human illness. Notable clinical manifestations include encephalitis (associated with viruses of the California serogroup), febrile illnesses (associated with viruses of the Bunyamwera, Group C and Simbu serogroups) and hemorrhagic fever (associated with the reassortant Ngari virus of the Bunyamwera serogroup) (Barrett & Shope, 2005; Nichol, 2001; Gerrard et al., 2004).

Phlebovirus genus

The genus *Phlebovirus* contains approximately 60 official and tentative virus species, with 8 members isolated from geographic locations spanning both hemispheres associated with human illness (Nichol et al, 2005). The majority of phleboviruses are vectored by phlebotomine sand flies; however, there are notable exceptions. For example, the important human and veterinary pathogen and type species of the genus, Rift Valley fever virus is known to be transmitted by *Aedes* species mosquitoes as well as by aerosolization. Also, Uukuniemi virus of the Uukuniemi group, linked to illness in sea birds, is associated with the tick *Ixodes ricinus.*

Rift Valley fever virus was first isolated in East Africa in 1930 during an epizootic that was characterized by high rates of mortality and abortion in infected sheep as well as acute febrile illness in humans (Daubney et al., 1931). This isolation underscores the historical importance of phleboviruses as both human and veterinary pathogens. Indeed, outbreaks of human and veterinary illnesses throughout history, occurring in areas across the Middle East, Africa and Europe, have been both contemporaneously associated and retrospectively linked with phlebovirus infections dating back to Napoleonic times (Eddy

& Peters, 1980; Hertig & Sabin, 1964; Verani and Nicolleti, 1995; Weekly epidemiological record, 2008). Of special interest, "sandfly fever" was responsible for significant troop morbidity during World War I and World War II, demonstrating the additional impact of phlebovirus infections on military history (Hertig & Sabin, 1964).

The majority of phlebovirus human infections are characterized by self-limiting febrile illnesses. Typical "sandfly fever" symptoms include the sudden onset of fever, malaise, anorexia, photophobia, abdominal symptoms and rash (Barrett & Shope, 2005). These symptoms are generally associated with both Old World (Naples and Sicilian) and New World (Alenquer, Candiru, Chagres, and Punta Toro) sand fly fever viruses. Similarly, Rift Valley fever virus human infections are most often associated with a self-limiting febrile illness characterized by the acute onset of fever along with abdominal symptoms (Barrett & Shope, 2005; Schmaljohn & Nichol, 2007). Unfortunately, a small subset of Rift Valley fever virus human cases can progress into hemorrhagic fever, hepatitis, encephalitis and/or retinal vasculitis (Barrett & Shope, 2005; Schmaljohn & Nichol, 2007) representing the most severe clinical manifestations associated with a phlebovirus infection.

Nairovirus genus

The genus *Nairovirus* is named for Nairobi sheep disease virus which was originally isolated in 1910 by the inoculation of healthy sheep with the blood of sheep with acute gastroenteritis in East Africa (Montgomery, 1917). There are seven identified species and 34 virus strains identified within the genus (Nichol et al, 2005). Nairoviruses are globally distributed and are primarily associated with tick vectors of transmission. However,

certain virus strains of the medically important Crimean-Congo hemorrhagic fever and Nairobi sheep disease groups have also been associated with culicoid flies and mosquitoes.

Only two virus strains of the genus *Nairovirus* have been identified as human pathogens, Nairobi sheep disease virus and Crimean-Congo hemorrhagic fever virus. Nairobi sheep disease virus has been linked with a mild, febrile illness in herdsmen that attend infected livestock during veterinary epidemics (Montgomery, 1917; Swanepoel, 1995). In clinical contrast, Crimean-Congo hemorrhagic fever virus causes severe human illness that is characterized by the early onset of flu-like symptoms followed by acute circulatory collapse, shock and hemorrhage resulting in death (Barrett & Shope, 2005). Cases of Crimean-Congo hemorrhagic fever have been documented in the Mediterranean, Africa, Eastern Europe, Central Asia, north western China and the Indian peninsula (Barrett & Shope, 2005). Interestingly, nosocomial infections of medical personnel and family members of infected individuals have been well documented (Burney et al., 1976). Also cases of Crimean-Congo hemorrhagic fever have been described in individuals involved in the slaughter of infected animals (Swanepoel et al., 1985). Taken together, these events implicate the aerosolization of infected blood products as an additional route of Crimean-Congo hemorrhagic fever virus transmission.

Physical properties of viruses of the family Bunyaviridae

Bunyaviruses are spherical, enveloped viruses of an approximate 100 nm in diameter (Figure 1.1). Glycoprotein projections, embedded within the lipid envelope, extend 5-10 nm from the surface of the virion (Figure 1.1). The viral particle contains a tripartite

genome of mostly negative polarity. The three genomic segments, L, M and S can occur in end-hydrogen bonded circularized forms and generally encode an RNA dependent RNA polymerase (L), envelope glycoproteins (Gn and Gc) and a nucleocapsid (N) protein, respectively. Each genomic segment is complexed within a ribonucleocapsid that contains both an abundance of N and a minority of L proteins (Figure 1.1). A unique property of the bunyavirus genome is the highly conserved, complementary nature of the 5' and 3' termini of each genomic segment, which facilitate circularization within the ribonucleocapsid (Table 1.2; Figure 1.1). Of additional interest, the organization of the coding regions within each genomic segment varies by genus, demonstrating a molecular basis for the extraordinary diversity of species described within the family (Figure 1.2).

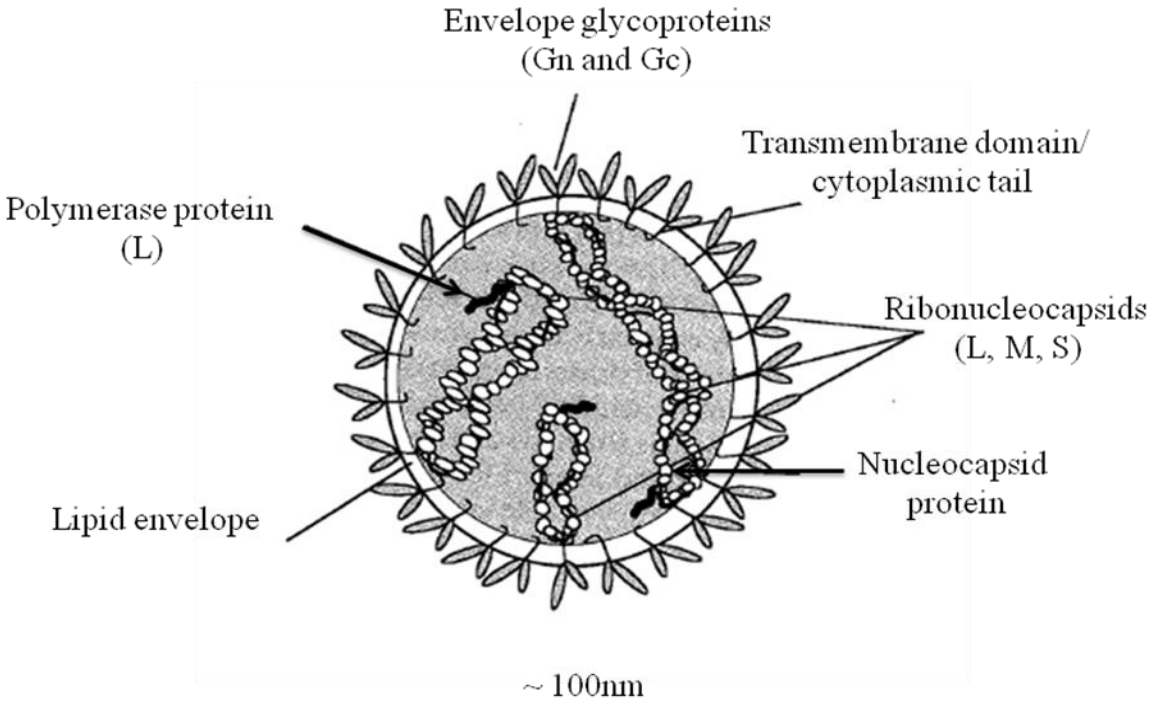

Figure 1.1 Generalized bunyavirus virion. Adapted from Schmaljohn & Nichol, 2007.

Table 1.2 Complementary terminal sequences
that are generally conserved among the S, M and L segments
of the family *Bunyaviridae.*

Genus	Sequence
Orthobunyavirus	3' UCAUCACAUG
	5' AGUAGUGUGC
Hantavirus	3' AUCAUCAUCUG
	5' UAGUAGUAUGC
Nairovirus	3' AGAGUUUCU
	5' UCUCAAAGA
Phlebovirus	3' UGUGUUUC
	5' ACACAAAG
Tospovirus	3' UCUCGUUA
	5' AGAGCAAU

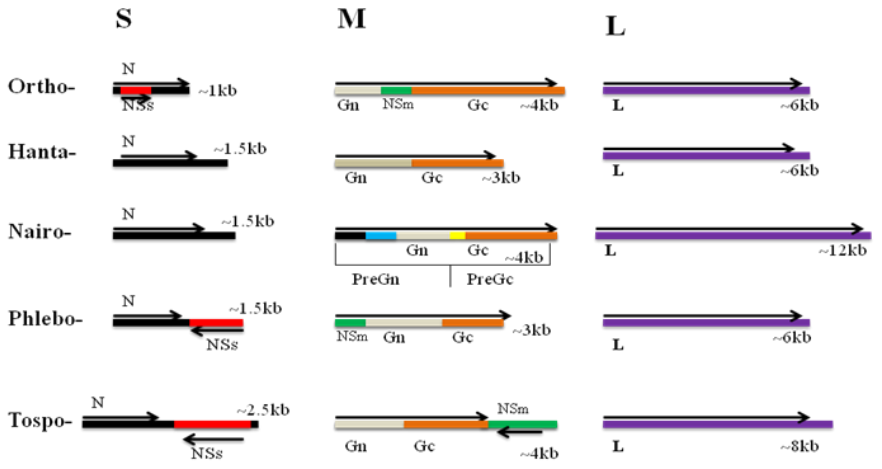

Figure 1.2. Generalized coding strategies of the family *Bunyaviridae*. Orthobunyavirus, hantavirus nairovirus, phlebovirus and tospovirus vcRNAs are depicted 5'-3'. ORF directionality is indicated by arrows and the relative sizes of ORFs and RNA segments are not drawn to scale. Of interest, the N and NSs overlapping ORFs of the S segment of the genus *Orthobunyavirus* and the ambisense nature of the phlebovirus and tospovirus S and M segments clearly represent the molecular diversity of the family *Bunyaviridae.*

Generalized replication strategy of Bunyaviridae

The critical functions of the key proteins encoded by the bunyavirus genome are well illustrated within the generalized replication strategy of *Bunyaviridae*, which is entirely cytoplasmic and begins with the interaction of virus surface glycoproteins with host cell receptors. Both Gn and Gc proteins have been implicated in host cell attachment (Arikawa et al., 1989; Keegan & Collett, 1986). However, Gc appears to be the primary attachment protein for members of the genus *Orthobunyavirus* (Plassmeyer et al., 2005). After attachment, uptake into the cell is driven by receptor-mediated endocytosis. Acidification of endocytic vesicles is thought to cause conformational changes in Gn and/or Gc, allowing fusion of viral and cellular membranes (Gonzalez-Scarano et al., 1984), thus facilitating the release of the ribonucleocapsid into the cytoplasm. Primary transcription of mRNA is primed by "cap snatching" from cytoplasmic host cellular mRNAs through endonucleolytic activity of the L protein (Patterson et al., 1984) and is facilitated by the N protein (Mir & Panganiban, 2005). The L protein is responsible for both primary transcription of mRNAs and genomic replication via a positive sense intermediate (Jin & Elliott, 1993). At some point in the replication cycle, L protein activity switches from primed, primary transcription to unprimed genomic replication through an unknown mechanism. For the M segment, translation of polypeptides occurs via ER bound ribosomes and the nascent polypeptide is co-translationally cleaved to generate Gn and Gc, which are dimerized within the ER (Schmaljohn & Nichol, 2007). Translation of L and S segment encoded proteins occurs on free ribosomes in the cytoplasm (Schmaljohn & Nichol, 2007). Virus assembly and maturation occurs in the ER and golgi membranes (Lyons and Heyduk, 1973; Novoa et al., 2005) and release of

mature virions occur by either transport through migration of golgi vesicles and fusion of vesicular and plasma membranes (Smith and Pafit, 1982; Bishop, 1996) or alternatively, via direct budding at the plasma membrane (Ravkov, 1997).

While not directly implicated within the generalized replication strategy, nonstructural proteins, NSs and NSm that are encoded within the *Orthobunyavirus, Phlebovirus* and *Tospovirus* genomes (Figure 1.2) are of growing research interest. Of these proteins, NSs functions of have been most well characterized in orthobunyaviruses as contributing to the shutdown of mammalian, but not mosquito host cell protein synthesis (Weber et al, 2001; Thomas et al, 2004; Hart et al, 2008). In addition, ortho- and phlebovirus NSs proteins have well documented anti-IFN activity in mammalian systems (Blakqori et al, 2007; Weber et al, 2002; Habjan et al, 2009). As such, NSs is thought to have a role in potentiating the zoonotic capacity of orthobunyaviruses (Hart et al, 2008) and phleboviruses by allowing these viruses to overcome vertebrate host innate immune responses.

Critical factors that have historically confounded the research and understanding of viruses of the family Bunyaviridae

i) *The advent of reverse-genetics systems for the investigation of negative-strand virus systems was historically delayed by the unique requirements of their replication strategies.* These requirements include the need for accessory proteins and intact 5′ and 3′ genomic ends for replication and packaging. Additionally, progress in the genetic manipulation of bunyaviruses has been further challenged by the

requirement of a multi-segment reverse-genetics system which is technically very labor-intensive. In recent years, researchers have made great progress in the advancement of molecular platforms for the genetic manipulation of negative-strand viruses despite the described challenges. Their efforts have generated functional systems for the molecular manipulation of representative bunyaviruses from multiple genera (Blakqori &Weber, 2005; Bridgen & Elliott, 1996; Flick & Pettersson, 2001; Ikegami et al, 2005); illustrating a break-through in the ability to investigate bunyavirus protein functions and host cell interactions. Notable among these systems are those developed within the laboratories of Richard Elliott and Friedemann Weber (Blakqori &Weber, 2005; Bridgen & Elliott, 1996) that have been utilized for the extensive study of protein functions for Bunyamwera and La Crosse viruses, respectively. Unfortunately, because of the labor intensive nature of their generation, these systems are likely to continue to be limited in number for just a handful of viruses among the hundreds classified within *Bunyaviridae*.

ii) *The large number and incredible diversity of viruses classified within the family Bunyaviridae have historically prevented their comprehensive molecular characterization.* At the initiation of the research described within this dissertation (2006), the vast majority of viruses within the family *Bunayviridae* entirely lacked molecular description. Among arthropod-borne human pathogens for which there

were available data in Genbank, the majority of sequences were partial in nature and limited to the S and/or M segments of the bunyavirus genome (Table 1.1). This lack of descriptive nucleotide sequence data, most certainly a function of the size and diversity of the family *Bunyaviridae*, has prevented the advanced understanding of *Bunyaviridae* at the molecular and evolutionary levels while also posing a great challenge to the detection of these viruses in the diagnostic and reference laboratory settings.

Evolution of Bunyaviridae and segment reassortment

The extraordinary diversity of the family *Bunyaviridae* bespeaks of both an ancient origin and an entirely unique evolutionary potential among all other known virus families. Indeed, for the arthropod-borne animal viruses of the family, the demanding ecology of their transmission cycle requires enormous plasticity for virus survival and propagation in nature. Presumably potentiating this capability, in addition to evolution through genetic drift, the segmented nature of the bunyavirus genome allows for the possibility of evolution through the reasortment of genomic segments between heterologous viruses. In fact, coinfection with more than one bunyavirus resulting in the reassortment of genomic segments has been well established in both laboratory and natural settings (Beaty et al., 1985; Beaty et al., 1997; Bishop & Beaty, 1988; Borucki et al., 1999; Briese et al., 2007; Cheng et al., 1999; Li et al., 1995; Nunes et al., 2005; Yanase et al., 2006). Of particular importance to analyses of the arthropod-borne genera, mosquitoes have been shown to be effective reservoirs for the reassortment of genomic

segments between heterologous bunyaviruses of shared serological character (Beaty et al., 1985; Beaty et al., 1997; Borucki et al., 1999). In addition, bunyavirus segment reassortment has been associated with outbreaks of human disease (Bowen et al., 2001; Briese et al., 2006; Gerrard et al., 2004). Despite this significant documentation, the role of segment reassortment in bunyavirus evolution and pathogenicity is largely unknown due to the previously described lack of comprehensive sequence data for members of the family *Bunyaviridae* (Gerrard et al., 2004).

Standard methods for the identification and characterization of arthropod-borne bunyaviruses

For diagnosis, surveillance and research the isolation of arthropod-borne bunyaviruses is performed through the inoculation of either suckling mice or susceptible cells (e.g. Vero cells) with sera or supernatants of homogenates derived from tissues of infected individuals or mosquito pools. Following isolation, identification and characterization of newly derived bunyavirus isolates has been historically provided through the use of predominantly antibody based methods (Lanciotti and Tsai, 2007). These methods include immunofluorescence assays (IFAs) that utilize grouping fluids of antibodies developed against multiple serologically related viruses, to classify an isolate at the serogroup level. Following serogroup level classification, plaque reduction neutralization tests (PRNTs), that use antibodies directed against individual species are then performed for species level identification. Together, these methods are irreplaceable for the identification of bunyaviruses in a group specific manner. However, they are time-consuming and can be limited in their ability to generate unequivocal, species identifying

results due to antibody cross reactivity and/or a limited diversity of available antibodies. In addition, because antibody based methods are limited to the detection of M segment encoded proteins, they are incapable of detecting reassortant viruses; a significant hinderence to the accurate and comprehensive detection of *Bunyaviridae.*

RT-PCR as an alternative to antibody-based methods for the identification and characterization of virus isolates

The advent of molecular technologies, such as RT-PCR, and the application of these technologies in diagnostic and reference laboratories, provides a time-efficient alternative to traditional serological methods for the identification of virus isolates. An attractive quality of nucleic acid based assays is the inherent flexibility of design which allows a researcher to develop an assay of varying degrees of specificity depending on their laboratory needs. Also, oligonucleotide primers that provide the fundamental binding events associated with the detection of viral nucleic acid can be readily synthesized according to the researcher's design. Additionally, the application of RT-PCR based molecular consensus assays, designed to detect a group of viruses of interest, followed by nucleotide sequencing for result confirmation and virus speciation has proven a powerful tool for emergent virus identification and discovery when applied to virus isolates in our laboratory (Lanciotti et al, 1999; 2007; 2008). However, molecular identification of any agent is only possible with *a priori* knowledge of species-identifying genomic sequence data, a limiting factor in the usefulness of nucleic acid based technologies for bunyavirus identification and characterization.

Central hypotheses and applied public health rationale that support the scope of research presented within this dissertation

It is believed that arthropod-borne viruses of the family *Bunyaviridae* will continue to be agents of public health import throughout the 21st century, as has been suggested by their frequent and devastating emergence in recent years (Bird et al., 2008; Gerrard et al., 2004; Vorou et al., 2007). Also, it is believed that the prevalence and medical relevance of viruses of the family *Bunyaviridae* are underestimated on a global scale because of their extreme diversity, highly limited characterization and segmented nature, which have historically precluded their rapid, accurate and comprehensive detection. At the Centers for Disease Control and Prevention, Division of Vector-Borne Diseases we have a specific mission to detect arthropod-borne human pathogens of the family *Bunyaviridae*, as they emerge. To support this mission, the following Specific Aims have been undertaken to address deficiencies in the comprehensive nucleotide sequence level characterization and detection of these viruses as well as to enhance the understanding of the evolution of the virus family *Bunyaviridae*.

Specific Aims and testable hypotheses that define the research presented within this dissertation

Specific Aim 1: Generate multi-segment primary nucleotide sequence data for arthropod-borne human pathogens and related viruses of the family *Bunyaviridae* that are not well described at the nucleotide sequence level.

Hypothesis: *Primary nucleotide sequence data can be generated for viruses of disparate classification within the family Bunyaviridae using amplification and sequencing methodologies that target highly conserved regions of the bunyavirus genome.*
Rationale: As previously described, a paucity of comprehensive nucleotide sequence data is highly problematic for the research, understanding and detection of viruses of the family *Bunyaviridae.* As such, a strategy for the generation of primary sequence data from viruses of unknown nucleotide sequence description is proposed. This strategy will exploit regions of known nucleotide sequence description and conservation among reference viruses of divergent classification within the family *Bunyaviridae*. It is assumed that these conserved regions, including the highly conserved termini of each genomic segment, are shared among viruses of known and unknown description and as such are ideal targets for our efforts.

Specific Aim 1 is addressed in part in Lambert and Lanciotti, 2008, Lambert and Lanciotti, 2009 and in Chapter 2 of this document.

Specific Aim 2: Develop a consensus assay for the detection of the majority of arthropod-borne human pathogens and related viruses of the family *Bunyaviridae.*

Hypothesis: *Despite extraordinary diversity, the majority of human pathogens within the family Bunyaviridae can be detected by a single nucleic acid based strategy.*

Rationale: As mentioned previously, RT-PCR based methodologies are time efficient and highly flexible alternatives to traditional serological methods for species and strain level identification of bunyaviruses of interest. Utilizing data generated in support of Specific Aim 1 (Chapter 2), we will compile a library of newly derived and previously generated nucleotide sequences as targets for the development of a broadly reactive consensus assay. This assay will be designed, developed and characterized for the detection of the majority of human pathogens within the *Orthobunyavirus, Nairovirus* and *Phlebovirus* genera of the family *Bunyaviridae.*

Specific Aim 2 is addressed in Lambert and Lanciotti, 2009 and in Chapter 3 of this document.

Specific Aim 3: Apply newly developed molecular methods for the detection of bunyaviruses of diverse origin in the diagnostic setting.

Hypothesis: *The application of newly developed molecular methods will facilitate the discovery of bunyaviruses in novel host species and geographic locations.*

Rationale: Based upon the belief that these viruses are likely underestimated in geographic distribution and host species range, we will apply molecular methods developed in support of Specific Aims 1 and 2 (Chapters 2 and 3) for the detection and identification of arthropod-borne bunyaviruses of unique circumstance. These methods will be applied to samples received in our international reference laboratory at the Centers for Disease Control and Prevention as part of our applied public health directive.

Specific Aim 3 is addressed in part in Kay et al, 2010; Lambert et al, 2010 and Lu et al, 2009 as well as in Chapter 4 of this document.

Specific Aim 4: Perform molecular and phylogenetic analyses on newly derived S segment nucleotide sequences to derive ancestral relationships among medically important members of the genus *Orthobunyavirus*.

Hypothesis: *The Bwamba serogroup is not a subgroup of the Bunyamwera serogroup of viruses.*

Rationale: Primary S segment nucleotide sequence data that were generated according to Specific Aim 1 (Chapter 2) include the first molecular description of the Bwamba serogroup viruses, Bwamba and Pongola viruses of Africa. These viruses were assumed by some to be related to viruses of the Bunyamwera serogroup of the genus *Orthobunyavirus* according to their generally shared clinical manifestations (febrile illnesses) and African geography (Yandodo et al, 2007). However, prior attempts at characterizing these agents at the molecular level utilizing methods that generally targeted the Bunyamwera serogroup failed to generate nucleotide sequence data (Yandoko et al, 2007). As such, the pursuit of molecular and phylogenetic analyses of these sequences is justified to facilitate an enhanced understanding of the molecular systematics of these and other medically important viruses of the genus *Orthobunyavirus.*

Specific Aim 4 is addressed in Lambert and Lanciotti, 2008 and in Chapter 5 of this document.

Specific Aim 5: Perform molecular and phylogenetic analyses on newly derived and previously determined nucleotide sequence data to evaluate the contribution of segment reassortment to the evolution of the California serogroup of the genus *Orthobunyavirus*.

Hypothesis: *S and M genomic segment reassortment among distinct viruses is not a central driving force in the evolution of the California serogroup of the genus Orthobunyavirus.*

Rationale: In the process of characterizing the molecular consensus assay developed according to Specific Aim and 2 (Chapter 3), we evaluated species and strain identifying S segment data for greater than 40 arthropod-borne bunyaviruses. Surprisingly, nearly all of the derived S segment nucleotide identities confirm the serologically inferred M segment identities of these viruses. These findings suggest a limited role for genomic segment reassortment in the evolution of many viruses of the genus *Orthobunyavirus*. These observations justify the evaluation of multi-segment data to examine the role of segment reasortment in the evolution human pathogens and related viruses of the genus *Orthobunyavirus*. In the interest of data management, our efforts will focus on analyses of the California serogroup for which serogroup level multi-segment analyses have not yet been performed.

Specific Aim 5 is addressed in Chapter 6 of this document.

CHAPTER 2

THE DETERMINATION OF PRIMARY NUCLEOTIDE SEQUENCE DATA FOR MULTIPLE GENOMIC SEGMENTS FROM BUNYAVIRUSES OF PREVIOUSLY UNDETERMINED OR INCOMPLETELY DESCRIBED NUCLEOTIDE SEQUENCE

Abstract

Primary nucleotide sequences have been determined for multiple genomic segments of medically relevant and related viruses of the arthropod-borne genera of the family *Bunyaviridae*. The amplification and sequencing strategy used to generate these sequences exploits regions of nucleotide sequence conservation shared between agents of disparate serological classification. The described efforts are fundamental to the development of nucleic acid based detection methods for these viruses and the advanced evolutionary understanding of the virus family *Bunyaviridae*.

Introduction

To support the detection and characterization efforts (Chapters 3, 4, 5, & 6), 47 full length and partial primary S, M and L nucleotide sequences of 31 viruses, including 14 human pathogens of the *Orthobunyavirus* and *Phlebovirus* genera of the family *Bunyaviridae* are presented (Table 2.1; Appendix 1). Viruses were selected for nucleotide sequencing after reviews of current literature and the NCBI GenBank database were conducted to identify arthropod-borne human pathogens and serologically related viruses of the family *Bunyaviridae* that lacked comprehensive molecular characterization at the beginning of this study, in 2006.

The work described in this chapter, ongoing since 2006, originally focused on the determination of S segment sequences across genera as a most tractable first priority to support the development of the presented molecular consensus assay (Chapter 3; Lambert & Lanciotti, 2009). In addition, M and L segment sequences have been determined to support the comprehensive molecular and evolutionary analyses described in Chapter 6 of this dissertation as well as those to be pursued in future studies. The overall goal of these efforts is to address deficiencies in molecular characterization so that the combination of newly determined and previously derived nucleotide sequence data that are available in GenBank will provide at least a minimal description of S, M and L genomic segments for viruses of interest (Table 2.1). Because of handling restrictions on human pathogens of the genus *Nairovirus*, the efforts detailed in this chapter do not include work on these agents.

Table 2.1. Human pathogens and serologically related viruses of the *Orthobunyavirus* and *Phlebovirus* genera for which primary nucleotide sequence data are presented[a]

Genus	Group	Virus	Strain	Source, Location and Date of Isolation	Segment amplified and sequenced in this study
Orthobunyavirus	Nyando	**Nyando**	MP 401	*Anopheles funestus; Kenya, 1959*	pt. L
	Guaroa	**Guaroa**	CoH 352111	*human; Colombia, 1956*	pt.L
	California	Trivittatus	Eklund	*Aedes trivittatus; North Dokota, USA, 1948*	pt.L
	California	Tahyna	Bardos 92	*Aedes caspius; Czechoslovakia, 1958*	pt.L
	California	South River	NJ0-94F	*Anopheles crucians; New Jersey, U.S.A.1966*	pt.L
	California	Jerry Slough	BFS 4474	*Culiseta inorata; California, U.S.A. 1963*	pt.L
	California	Serra do Navio	BeAr 103645	*Aedes fulvus; Brazil, 1966*	pt.L
	California	**California Enc.**	85-415NM	*Aedes campestris; New Mexico, USA, 1985*	pt.L
	Bunyamwera	**Fort Sherman**	86MSP18	*human; Panama, 1985*	full S, pt. L
	Bunyamwera	**Shokwe**	Sa Ar 4042	*Aedes cumminsii; South Africa, 1962*	full S, pt. L
	Bunyamwera	**Wyeomyia**	Original	*Wyeomyia melanocephala; Colombia, 1940*	full S, pt. L
	Bunyamwera	**Xingu**	BeH 388464	*human; Brazil, unknown date*	full S, pt. L
	Bunyamwera	Birao	DakArB 2198	*Anopheles pharoensis; Central African Rep., 1969*	pt. L
	Bunyamwera	Bozo	DakArB 7343	*Aedes spp; Central African Rep., 1975*	pt. L
	Bunyamwera	Lokern	FMS 4332	*Culex tarsalis; California, USA, 1962*	pt. S, pt. M, pt. L
	Bunyamwera	Playas	75V3066	*Aedes taeniorhynchus; Playas, Ecquador, 1975*	pt. S, pt. M, pt. L
	Bunyamwera	Iaco	BeAr 314206	*Wyeomyia spp; Brazil, 1976*	pt. L
	Bunyamwera	Santa Rosa	M2-1493	*Aedes angustivittatus; Mexico, 1972*	pt. S, pt. M, pt. L
	Bwamba	**Bwamba**	M 459	*human; Uganda, 1937*	full S, pt. L
	Bwamba	**Pongola**	Sa Ar 1	*Aedes (N.) circumluteolus; South Africa, 1955*	full S, pt. L
	Catu	**Catu**	BeH 151	*human; Brazil, 1955*	pt. L
	Guama	**Guama**	BeAn 277	*Cebus apella; Brazil, 1955*	pt. L
Phlebovirus	Sandfly fever group	**Alenquer**	BeH 301101	*human; Brazil, 1976*	pt. S, pt. M, pt. L
	Sandfly fever group	**Candiru**	BeH 22511	*human; Brazil, 1960*	pt. S, pt. M
	Sandfly fever group	**Chagres**	JW 10	*human; Panama, 1960*	pt. S, pt. L
	Sandfly fever group	Cacao	VP 437R	*Lutzomyia trapidoi; Panama, 1970*	pt. S
	Sandfly fever group	Itaporanga	Original	*sentinal mouse; Brazil, 1962*	pt. S
	Sandfly fever group	Gabek Forest	SudAn 754-61	*Acomys albigena; Sudan, 1961*	pt. S
	Sandfly fever group	Turuna	Bear352492	*Lutzomyia spp; Brazil, 1978*	pt. S
	Sandfly fever group	Rio Grande	TBM3-204	*Neotoma micropus; Texas, U.S.A., 1973*	pt. S
	Sandfly fever group	Nique	9c	*Lutzomyia panamensis; Panama, 1972*	pt. S

[a] *Human pathogens are emboldened*

Materials and Methods

All described procedures were performed according to manufacturers' instructions unless stated otherwise.

Viruses. Viruses were provided by the Arbovirus Diseases Branch of the Division of Vector-Borne Diseases, National Center for Zoonotic, Vector-Borne, and Enteric Diseases (NCZVED), Centers for Disease Control and Prevention (CDC), World Health Organization (WHO) arthropod-borne virus reference collection. Low passage preparations of selected virus strains, of a minimum titer of 10000 PFU/mL in Vero cells, from the *Orthobunyavirus* and *Phlebovirus* genera were used in this study (Table 2.1).

RNA extraction. Viral RNA was extracted from virus seed using the QIAamp Viral RNA mini kit (Qiagen). Extractions were performed on samples ranging in volume from 70-140 µl. RNA was eluted in a volume equal to the volume of starting sample at a

minimum concentration of 100 PFU/µl. Eluted RNA was stored at -70ºC until evaluated by RT-PCR amplification.

Primer design and evaluation. For cDNA amplification and nucleotide sequencing from viruses of interest, more than 250 oligonucleotide primers were designed to target S, M and L segment sequences of the *Orthobunyavirus* and *Phlebovirus* genera of the family *Bunyaviridae*. To facilitate primer design, alignments were made from available full length and partial S, M and L sequences (Table 2.2) that were selected as representative of the serological diversity of viruses classified within each genus. Sequences were aligned by genus and serogroup, using the MegAlign software (DNA*) Clustal W alignment function (Thompson, 1994). Primers were then manually designed according to the visual identification of regions of relative conservation. Newly generated primers were then evaluated by RT-PCR-based application to RNAs extracted from genus-appropriate reference viruses (Table 2.2). According to gel-based detection and nucleotide sequence analysis, oligonucleotide primers that amplified target cDNAs from the broadest diversity of reference virus RNAs (Table 2.2) were chosen (Table 2.3) for evaluation against bunyaviruses of interest (Table 2.1). Regions that are targeted by these primers include the highly conserved 5’ and 3’ termini of the orthobunyavirus and phlebovirus S, M, and L segments and are presented in Figures 2.1-2.6.

Table 2.2. Sequences of human pathogens of the *Orthobunyavirus* and *Phlebovirus* genera used as references for consensus primer design				
Genus		**GenBank #**	**GenBank #**	**GenBank#**
Orthobunyavirus	**Virus**	**S seg. sequence**	**M seg. sequence**	**L seg. sequence**
Bunyamwera	Bunyamwera	NC_001927	NC_001926	NC_001925
Bunyamwera	Cache Valley	X73465		
Bunyamwera	Garissa (Ngari)	AF398345		
Bunyamwera	Germiston	M19420		
Bunyamwera	ilesha	AY729651		
Group C	Apeu	DQ188952		
Group C	Caraparu	DQ188974		
Group C	Itaqui	DQ188984		
Group C	Madrid	DQ188957		
Group C	Marituba	DQ188980		
Group C	Murutucu	DQ188972		
Group C	Nepuyo	DQ188981		
Group C	Oriboca	DQ188967		
Group C	Ossa	DQ188954		
Group C	Restan	DQ188956		
California	CE	U12800		
California	Guaroa	X73466		
California	Inkoo	Z68496		
California	La Crosse	NC_004110	NC_004109	NC_004108
California	SSH	J02390		
California	Tahyna	Z68497		
Nyando	Nyando	AM709781		
Simbu	Oropouche	NC_005777	NC_005775	NC_005776
Genus				
Phlebovirus				
Uukuniemi	UUKV			NC_005214
Sandfly fever group	Massilia		EU725772	
Sandfly fever group	Punta Toro	DQ363406	DQ363407	
Sandfly fever group	RVF	NC_002045	NC_002044	NC_002043
Sandfly fever group	SFN	EF201832		
Sandfly fever group	Toscana	EF201833	NC_006320	NC_006319
Sandfly fever group	SFS	AJ811547		
CE, California Encephalitis virus; SSH, Snowshoe hare virus; UUKV, Uukuniemi virus				

	5’ Terminus, NCR	3’ Terminus, NCR
La Crosse virus	5’ [1]AGUAGUGUACUCCAC……..	..GUGGAGCACACUACU [984] 3’
Inkoo virus	5’ [1]AGUAGUGUACUCCAC……..	..GUGGAGCACACUACU [986] 3’
Bunyamwera virus	5’ [1]AGUAGUGUACUCCAC……..	..GUGGAGCACACUACU [961] 3’
Cache Valley virus	5’ [1]AGUAGUGUACUCCAC……..	..GUGGAGCACACUACU [950] 3’
Apeu virus	5’ [1]AGUAGUGUACUCCAC……..	..GUGGAGCACACUACU [916] 3’
Oropouche virus	5’ [1]AGUAGUGUACUCCAC……..	..GUGGAGCACACUACU [754] 3’

Figure 2.1 Conserved regions of selected orthobunyavirus S segment vcRNA 5’ and 3’ termini that are targeted by the presented amplification and sequencing strategy. Nucleotide position designations for each virus sequence appear at the 5’ and 3’ ends of presented sequences. Conserved regions are highlighted.

	5’ Terminus, NCR	H----A----C-----F-----
La Crosse virus	5’ [1]AGUAGUGUACUACC……..	UCCAUGCCUGCUUCAA [585] 3’
Bunyamwera virus	5’ [1]AGUAGUGUACUACC……..	UCCAUGCUUGCUUUAA [589] 3’
Oropouche virus	5’ [1]AGUAGUGUACUACC……..	UCCAUGCAUGCUUUAG [579] 3’

Figure 2.2 Conserved regions of selected orthobunyavirus M segment vcRNAs that represent regions

that are targeted by the presented amplification and sequencing strategy.

Nucleotide position designations appear at the 5' and 3' ends of presented sequences.
The 5' and 3' targeted regions occur at the 5' terminus and within the
Gn portion of the orthobunyavirus polyprotein ORF, respectively. For the Gn target,
corresponding amino acids appear above the nucleotide sequence.
Regions of conservation are highlighted.

	5’ terminus, NCR	--D----D-----E/D--E/K-
La Crosse virus	5’ [1]AGUAGUGUACCCCUA……..	UGAUGAUGAAGAAUU [594] 3’
Bunyamwera virus	5’ [1]AGUAGUGUACUCCUA……..	UGAUGAUGAGGAAUU [583] 3’
Oropouche virus	5’ [1]AGUAGUGUGCUCCUA……..	GGAUGAUGAUAAAUU [579] 3’

Figure 2.3 Conserved regions of selected orthobunyavirus L segment vcRNAs that

represent regions that are targeted by the presented amplification and sequencing strategy.

Nucleotide position designations for each virus sequence appear at the 5’ and 3’ ends of presented sequences.
The 5' and 3' targeted regions occur at the 5' terminus of the L segment and within the
the orthobunyavirus polymerase ORF, respectively. For the polymerase target, corresponding amino
acids appear above the nucleotide sequence.
Conserved regions are highlighted.

Punta Toro virus

F----A----Y---Q---G----F----D---- M----M---H---P----S----F----A---G----V/L-I/V-

5’[113]UUUGCUUAUCAAGGAUUUGAUGC..... AUGAUGCAUCCCAGCUUUGCUGGAGUGAUUGA[492]3’

Sandfly Fever Sicilian

5’[123] UUUGCGUAUCAAGGAUUUGAUGC.....AUGAUGCAUCCAAGCUUUGCCGGACUGAUUGA[505]3’

Rift Valley Fever

5’[82] UUUGCUUAUCAAGGGUUUGATGC……..AUGAUGCAUCCCAGCUUUGCUGGAGUGGUGGA[461]3’

Toscana

5’[101] UUUGCUUACCAAGGGUUUGAUGC……..AUGAUGCACCCCAGCUUUGCUGGAUUGAUUGA[492]3’

Sandfly Fever Naples

5’[85] UUUGCUUAUCAAGGAUUUGAUGC……..AUGAUGCAUCCUAGCUUUGCAGGACUGAUUGA[476]3’

Figure 2.4 Conserved regions of the phlebovirus N ORF (S segment) that are targeted by the presented amplification and sequencing strategy shown for selected viruses. N ORF nucleotide position designations for each virus sequence appear at the 5’ and 3’ ends of presented sequences. Corresponding amino acids appear above the nucleotide sequence

Conserved regions are highlighted.

	N----Y-----Q----C----	E-----N-----K----C----F-----E----
Massilia virus	5’ [2143]AAUUAUCAGUGCCA……	GAAAAUAAAUGCUUUGAGCA [2867]3’
Punta Toro virus	5’ [2098]AACUAUCAGUGCCA……	GAAAAUAAAUGCUUUGAACA [2810] 3’
Toscana virus	5’ [2158]AAUUACCAGUGCCA……	GAAAAUAGAUGCUUUGAGCA [2880] 3’

Figure 2.5 Conserved regions of selected phlebovirus M segment vcRNAs that represent regions that are targeted by the presented amplification and sequencing strategy.

Nucleotide position designations for each virus sequence appear at the 5’ and 3’ ends of presented sequences. Targeted regions span the Gn and Gc portions of the polyprotein ORF. Corresponding amino acid sequences appear above the nucleotide sequences.

Conserved regions are highlighted.

	K---K----Q----Q---H----G---G----	M----M---Q---G----I-----L---H---Y/F--T----S----S---
RFV virus	5’ [2770] AAGAAACAACAGCAUGGAGGU……	AUGAUGCAGGGAAUACUGCAUUAUACUUCCUCA [3300]3’
Toscana virus	5’ [2770]UUCAAGAAACAACAGCAUGGUGGU……	AUGAUGCAGGGGAUUCUGCACUUCACUUCCUCC[3303]3’
UUKV virus	5’ [2774]UUCAAGAAACCGCAGCAUGGUGGU……	AUGAUGCAAGGCAUACUUCAUUACACCUCUAGU [3330] 3’

Figure 2.6 Conserved regions of selected phlebovirus L segment vcRNAs that represent regions that are targeted by the presented amplification and sequencing strategy.

Nucleotide position designations for each virus sequence appear at the 5’ and 3’ ends of presented sequences. The 5' and 3' targeted regions occur within the polymerase ORF. Conserved regions are highlighted

.

Amplification of S, M and L segment cDNAs for primary nucleotide sequence analysis. RNAs extracted from selected viruses (Table 2.1) were subjected to RT-PCR using genus-appropriate primers (Table 2.3). Fifty pmol of each forward and corresponding reverse primer (Table 2.3) and 5µl of viral RNA, in a 6 µl total volume, were heated to 65 ºC for five minutes and immediately placed

Table. 2.3 Oligonucleotide primers selected for consensus amplification and nucleotide sequencing of cDNAs amplified from human pathogens and related viruses of the *Orthobunyavirus* and *Phlebovirus* genera.				
				amplicon size
1) Ortho S forward	AGTAGTGTRCTCCAC	**Ortho S reverse**	AGTAGTGTRCTCCAC	~ 1000 bp
2) Ortho M 5' forward	AGTAGTGTACTACC	**Ortho M reverse**	TTRAARCADGCATGGAA	~520 bp
3) Ortho L 5' forward	AGTAGTGTACTCCTA	**Ortho L reverse**	AATTCYTCATCATCA	~520 bp
4) Phlebo S forward 1	TTTGCTTATCAAGGATTTGATGC	**Phlebo S reverse**	TCAATCAGTCCAGCAAAGCTGGGATGCATCAT	~400 bp
Phlebo S forward 2	TTTGCTTATCAAGGATTTGACC			
5) Phlebo M forward	AAYTAYCAGTGYCA	**Phlebo M reverse**	TGYTCAAAGCATTTATTTTC	~600 bp
6) Phlebo L forward	TTCAAGAAAACARCAGC	**Phlebo L reverse**	AGAATTCCTTGCATCATCTTAAGGAACGTAGTA	~230 bp

on ice. Forty five µl of OneStep RT-PCR mastermix (Qiagen) was added to the RNA/primer mixture. Reactions were subjected to RT-PCR amplification using the following cycling conditions: one incubation of 45ºC for 60 minutes (RT), 95 ºC for 15 minutes (denaturation) ; followed by 55 cycles of 94 ºC for 30 seconds, 45ºC for 1 minute and 72 ºC for 2 minutes (PCR cycling). Reactions were terminated after a final extension step at 72 ºC for 10 minutes.

Visualization of DNA and DNA purification. Products of amplification were evaluated by electrophoresis in a 1.5% agarose gel in 40 mM Tris-Acetate-1mM EDTA (TAE) buffer. Visualization of DNA bands was achieved through ethidium bromide staining and UV transillumination. DNA bands of target size (~230-1000bps; Table 2.3) were excised from the agarose gel and purified using the QIAquick gel extraction kit (Qiagen).

Preparation of plasmid ligated *Orthobunyavirus* full length S segment cDNAs for primary nucleotide sequence determination. For *Orthobunyavirus* S segment sequence determination, purified full length cDNAs were ligated into the pCR2.1-TOPO vector (Invitrogen). Ligated plasmids were introduced to TOP 10F' competent *E. coli*

cells (Invitrogen) through chemical transformation. Transformed cells were spread onto plates containing imMedia Amp Blue agar (Invitrogen) and plates were incubated overnight at 37ºC. White colonies were picked, and 3mL of imMedia (Invitrogen) were inoculated with each white colony to be evaluated. Inoculated media were incubated, shaking, at 37ºC overnight. Plasmid DNAs were then extracted from 1.7mL of these *E. coli* preparations using the Qiagen miniprep kit (Qiagen).

To verify the presence, and determine the sequence of DNA inserts, ~100ng of each extracted plasmid DNA was sequenced with M13 forward and reverse primers provided with the pCR2.1-TOPO vector (Invitrogen) in individual reactions using ABI BigDye Terminator V3.1 Ready Reaction Cycle Sequencing mix (Applied Biosystems). Sequencing reactions were purified using the Qiagen DyEx 2.0 spin kit (Qiagen). Nucleotide sequences were determined by running purified sequencing reactions on the ABI 3130 genetic analyzer (Applied Biosystems). No fewer than three plasmid preparations, originating from three separate colonies, were sequenced for each virus. All sequences were confirmed by direct sequencing of purified RT-PCR products using virus specific primers designed from plasmid derived sequence. Terminal sequences were determined through 3' poly-(A) tailing of purified viral RNAs (Ambion) and 5'3' RACE (Roche) followed by nucleotide sequencing. Through the described methods, all presented full length S segment sequences were confirmed from multiple preparations of viral RNAs and RT-PCR amplifications, in both the 5' and 3'directions.

Nucleotide sequencing of RT-PCR amplified partial orthobunyavirus and phlebovirus cDNAs for primary nucleotide sequence determination. To determine the nucleotide sequences of purified cDNAs amplified from partial M and L segment viral RNAs of the genus *Orthobunyavirus* and S, M and L segment viral RNAs of the genus *Phlebovirus*, the following steps were followed. In independent reactions, approximately 20 ng of each cDNA was sequenced with 3.2 pmol of the appropriate forward primer and the corresponding reverse primer (Table 2.3) that were utilized in the amplification reaction by the methods described above. As with the presented full length sequences, all sequences were confirmed from multiple RNA preparations, amplification and sequencing reactions.

Analyses of newly generated sequences and submission of primary nucleotide sequence data to GenBank. Nucleotide sequences generated from all evaluated bunyaviruses were aligned and edited using SeqMan software (DNA*). Resultant consensus sequences were additionally characterized by the identification of complete and partial ORFs (for partial sequences) followed by translation of these regions using EditSeq software (DNA*). All nucleotide and amino acid sequences were then subjected to NCBI BLAST analyses (Altschul et al, 1990) to verify their identities at the serogroup and genus levels. All newly derived primary sequence data have been or will be submitted to GenBank for public access in accordance with their publication (Lambert and Lanciotti, 2008; Lambert and Lanciotti, 2009; Table 2.1; Appendix 1).

Results and Discussion

Through the above described methods, full and partial length S, M and L segment sequences were determined for viruses of the *Orthobunyavirus* and *Phlebovirus* genera of the family *Bunyaviridae* that previously lacked molecular descriptions in targeted

regions. Newly determined sequences include the premiere molecular description of the medically important Bwamba, Catu and Guama serogroups of the genus *Orthobunyavirus* and New World human pathogens Alenquer, Candiru and Chagres viruses of the genus *Phlebovirus* (Table 2.1; Appendix 1) among others.

In addition to virus genome segments for which primary nucleotide sequences are reported (Table 2.1), the amplification and sequencing of cDNAs from other virus segments have been attempted without success. Most notably, these efforts include multiple attempts utilizing the above described methods in addition to virus purification and attempted enzymatic selection of target viral RNAs (methods not described here) to enrich target sequence and enhance the likelihood of amplification of M and L segment cDNAs from Group C viruses of the genus *Orthobunyavirus.* The inability to amplify M and L segment cDNAs from Group C viruses, combined with the lack of sequence data in the literature and public databases, suggests that target nucleic acid heterogeneity exists among variant strains of these viruses. Interestingly, as discussed in Chapter 3, methods designed to detect Group C virus S segment RNAs, for which nucleotide sequence data exist (Nunes et al, 2005) similarly failed to generate cDNAs for sequencing, further suggesting variant nucleic acid character for divergent Group C strains. Future research, using alternative methods to elucidate sequences for Group C viruses, is supported by these observations. Of special interest is the investigation of potential heterogeneity at the termini of the S, M and L segments, that when targeted by the presented amplification strategies, designed to detect highly conserved sequences of the *Orthobunyavirus* genus, failed to generate Group C virus cDNAs for sequencing.

Of note, the described analyses of the presented nucleotide sequences have generated interesting findings that are not presented elsewhere within this dissertation. These

findings include the discovery of a truncated NSs ORF within the Wyeomyia virus S segment (see Appendix 1). This finding, combined with the understanding that NSs functions likely potentiate vertebrate host pathogenesis (Hart et al, 2009) makes this virus of unique research interest. Also of interest, preliminary NCBI BLAST analyses (Altschul et al, 1990) of New World phlebovirus L segment nucleotide sequences reveal a strong similarity with L segment sequences of the tenuiviruses, which are negative sense, segmented RNA plant viruses of limited description. This finding suggests a potential shared ancestry between the *Bunyaviridae* and tenuiviruses that has yet to be investigated.

The key goals for the determination of the presented bunyavirus sequences (Table 2.1; Appendix 1) are to 1) provide target sequences for the comprehensive detection and identification of these agents (Chapter 3; Lambert & Lanciotti, 2009) and 2) support the presented molecular characterization and evolutionary analyses (Lambert & Lanciotti, 2008; Chapters 5 & 6). Furthermore, it is hoped that the described efforts will also support the future research and understanding of the virus family *Bunyaviridae* as well.

CHAPTER 3

A CONSENSUS AMPLIFICATION AND NOVEL MULTIPLEX SEQUENCING METHOD FOR S SEGMENT SPECIES IDENTIFICATION OF 47 VIRUSES OF THE *ORTHOBUNYAVIRUS*, *PHLEBOVIRUS* and *NAIROVIRUS* GENERA OF THE FAMILY BUNYAVIRIDAE

Abstract

An RT-PCR assay was designed according to previously determined and newly derived genomic sequences to target S genomic segments of 47 viruses, including 29 arthropod-borne human pathogens, of the family *Bunyaviridae*. The analytical sensitivity of the presented assay was evaluated through its application to RNAs extracted from quantitated dilutions of bunyaviruses of interest. Additionally, the assay's analytical specificity was determined through the evaluation of RNAs extracted from selected bunyaviruses, and other representative arthropod-borne viruses, isolated from a diversity of host species, temporal, and geographic origins. After RT-PCR amplification, cDNAs from bunyaviruses of interest were subjected to a novel multiplex sequencing method to confirm bunyavirus positivity and provide preliminary, species level S segment identification. It is our goal that this assay will be used as a tool for identification and characterization of emergent arthropod-borne bunyavirus isolates of medical importance as well as related viruses of the family *Bunyaviridae* that have not been associated with human illness.

Introduction

Virus isolation followed by identification and characterization of the resultant isolate by available methods is the gold standard for the detection of arthropod-borne viruses of the family *Bunyaviridae*. Historically, predominantly serological methods have been used for the identification and characterization of bunyavirus isolates (Lanciotti & Tsai, 2007) because of their broad reactivity and unrivaled ability to identify bunyaviruses at the serogroup level. However, as previously discussed (Chapter 1), these methods are restricted in their ability to provide unequivocal, species level identification because of antibody cross reactivity and/or a limited diversity of available antibodies.

RT-PCR-based methods are time-efficient, highly sensitive and specific alternatives to traditional serological methods for the identification of virus isolates. However, the broad-scale molecular identification of arthropod-borne viruses of the family *Bunyaviridae* has been historically prevented due to limited nucleotide sequence data availability for many of these agents and the high level of diversity among bunyaviruses of determined genomic sequences (Chapter 1). Despite these significant challenges it is proposed that the majority of arthropod-borne human pathogens within the family *Bunyaviridae* can be identified and characterized by a single RT-PCR based assay.

To evaluate this assertion, both full and partial length S segment primary nucleotide sequence data have been determined for fifteen bunyaviruses that had previously not been determined as targets for assay design and detection prior to the development of the presented assay (Lambert & Lanciotti, 2008; Chapter 2). The Genbank/EMBL/DDBJ accession numbers for these newly determined S segment sequences are EU564827 (Bwamba virus), EU564828 (Pongola virus), EU564829 (Fort Sherman virus), EU564830 (Xingu virus), EU564831 (Shokwe virus), FJ235921 (Wyeomyia virus),

FJ235922 (Alenquer virus), FJ235923 (Chandiru virus), FJ235924 (Chagres virus), FJ235925 (Cacao virus), FJ235926 (Itaporanga virus), FJ235927 (Gabek Forest virus), FJ235928 (Turuna virus), FJ235929 (Rio Grande virus), FJ235930 (Nique virus) and GQ166188 (Itaituba virus). Here, the development of an RT-PCR based molecular consensus assay, incorporating a diversity of oligonucleotide primers, for the detection of 47 viruses of the *Orthobunyavirus*, *Nairovirus* and *Phlebovirus* genera of the family *Bunyaviridae*, including 29 human pathogens with newly derived and previously determined S segment sequences is described. Because of the high degree of characterization across genera, and the relative conservation within genera, the S segment N ORF is the genomic target of the presented assay.

The format of the presented RT-PCR-based molecular consensus assay includes the application of two consensus mixes of oligonucleotides for the detection of RNAs extracted from target viruses of i) the *Orthobunyavirus* and ii) the *Nairovirus* and *Phlebovirus* genera of the family *Bunyaviridae.* For S segment species determination, RT-PCR amplified bunyaviral cDNAs were subjected to a novel multiplex sequencing method followed by NCBI BLAST analyses of newly derived nucleotide sequence data. Based on S segment species determination, this assay is being presented as a tool for the preliminary identification of medically important, and serologically related, viruses of the *Orthobunyavirus*, *Nairovirus* and *Phlebovirus* genera of the family *Bunyaviridae.* While there are existing molecular consensus assays for the detection of serogroups of viruses within the family *Bunyaviridae* (Kuno et al, 1996; Sánchez-Seco et al, 2003), to the best of our knowledge, the presented assay detects a larger diversity of bunyaviruses than any previously published effort. It is our hope that this assay be used as a complement or

alternative to standard antibody-based methods for the identification and characterization of bunyaviruses isolated from a variety of geographic locations and sources.

Materials and Methods

All procedures were performed according to manufacturer's instructions unless stated otherwise.

Viruses. Viruses used in this study were provided by the Arbovirus Diseases Branch of the Division of Vector-Borne Infectious Diseases, National Center for Zoonotic, Vector-Borne, and Enteric Diseases (NCZVED), Centers for Disease Control and Prevention (CDC), World Health Organization (WHO) arthropod-borne virus reference collection. Viruses selected for evaluation include human pathogens and serologically related viruses of the family *Bunyaviridae* (Tables 3. 1 & 3.2). Viruses were identified at the species level through serological analyses prior to the beginning of this study. In a few cases, regulations for handling live viruses prevented our ability to work with these agents in our laboratory. As alternatives, we obtained Nairobi sheep disease RNA and Crimean-Congo hemorrhagic fever (CCHF) virus plasmid derived RT-PCR controls from George Ludwig (USAMRIID) and Rosemary Sang (KEMRI), respectively. For quantitation, human pathogens (Table 3. 1) were subjected to plaque titration of 10-fold dilutions in Vero cells prior to molecular evaluation. Additional viruses of the family *Bunyaviridae*, as well as representative arthropod-borne viruses of the *Alphavirus* and *Flavivirus* genera, of the virus families *Togaviridae* and *Flaviviridae*, were evaluated to determine assay specificity (Tables 3. 1 & 3.2).

Table 3.1. Bunyaviruses that are targeted and amplified by the molecular consensus assay[a]

Genus/ Serogroup[b]	Virus[c]	Strain	source, location and date of isolation	RT-PCR result / serological identity confirmed by nucleotide sequencing and BLAST analysis	Detection Limit (PFU/.1mL)[c]	
Orthobunyavirus						
Bunyamwera	**Bunyamwera**	Original	*Aedes spp.;* Uganda, 1943	**POS/YES**	**70PFU**	Orthobunyavirus primer mix
Bunyamwera	**Cache Valley**	6V633	*Culiseta inornata;* Utah, U.S.A.,1956	**POS/YES**	**600PFU**	
Bunyamwera	**Fort Sherman***	86MSP18	human; Panama, 1985	**POS/YES**	**100PFU**	
Bunyamwera	**ilesha**	KO/2	human; Nigeria, 1957	**POS/NO**	**300PFU**	
Bunyamwera	**Shokwe***	Sa Ar 4042	*Aedes cumminsii;* South Africa, 1962	**POS/YES**	**75PFU**	
Bunyamwera	**Wyeomyia***	Original	*Wyeomyia melanocephala;* Colombia, 1940	**POS/YES**	**660PFU**	
Bunyamwera	**Xingu***	BeH 388464	human; Brazil, unknown date	**POS/YES**	**4PFU**	
Bunyamwera	**Garissa (Ngari)**[d]	not evaluated		ND	ND	
Bunyamwera	**Germiston**	not evaluated		ND	ND	
Bunyamwera	Maguari	75V3429	*Aedes scapularis* ; Ecuador, 1974	**POS/NO**	ND	
Bunyamwera	Potosi	89-3380	*Aedes albopictus* ; Missouri, 1989	**POS/YES**	ND	
Bunyamwera	Northway	234	*Aedes spp* .; Alaska, 1971	**POS/YES**	ND	
Bunyamwera	Biroa	DakArB 2198	*Anopheles pharoensis* ; CAR, 1969	**POS/YES**	ND	
Bunyamwera	Batai	184	*Anopheles maculipennis;* Czechoslovokia, 1950	**POS/YES**	ND	
Bwamba	**Bwamba***	M 459	human; Uganda, 1937	**POS/YES**	**200PFU**	
Bwamba	**Pongola***	Sa Ar 1	*Aedes (N.) circumluteolus;* South Africa, 1955	**POS/YES**	**200PFU**	
Simbu	**Oropouche**	TRVL 9760	human; Tobago, 1955	**POS/YES**	**.1PFU**	
California	**California Encephalitis**	BFS 283	*Aedes melanimon;* California, U.S.A., 1944	**POS/YES**	**6PFU**	
California	**Inkoo**	KN 3641	*Aedes communis punctor;* Finland, unknown date	**POS/YES**	**1.3PFU**	
California	**Jamestown Canyon**	61V2235	*Culiseta inornata;* Colorado, U.S.A., 1967	**POS/YES**	**60PFU**	
California	**La Crosse**	Original	human; Wisconsin, U.S.A., 1964	**POS/YES**	**200PFU**	
California	**Snowshoe hare**	LEIV 13004AKH	*Aedes communis;* U.S.S.R., 1986	**POS/NO**	**34PFU**	
California	**Tahyna**	Bardos 92	*Aedes caspius;* Czechoslovakia, 1958	**POS/YES**	**340PFU**	
California	Serra de Navio	BeAr 103645	*Aedes spp.* ; Brazil, 1966	**POS/YES**	ND	
California	South River	NJ0-94F	*Anopheles crucians* ; New Jersey, U.S.A.1966	**POS/YES**	ND	
California	Jerry Slough	BFS 4474	*Culiseta inorata* ; California, U.S.A. 1963	**POS/YES**	ND	
Nairovirus						
CCHF	**Crimean-Congo HF**	S segment cDNA	plasmid derived PCR control	**POS/YES**	ND	Phlebo/Nairovirus primer mix
CCHF	Hazara	PakJC 280	*Ixodes redikorzevi;* Pakistan, 1964	**POS/YES**	ND	
NSD	**Dugbe**	IbAr 1792	*Amblyomma variegatum;* Nigeria, 1964	**POS/YES**	**2PFU**	
NSD	**Nairobi sheep disease**	RV082	RNA control from Kenyan strain	**POS/YES**	ND	
Phlebovirus						
SF	**Alenquer***	BeH 301101	human; Brazil, 1976	**POS/YES**	**0.02PFU**	
SF	**Candiru***	BeH 22511	human; Brazil, 1960	**POS/YES**	**0.3PFU**	
SF	**Chagres***	JW 10	human; Panama, 1960	**POS/YES**	**0.0024PFU**	
SF	**Punta Toro**	D 4021A	human; Panama, 1966	**POS/YES**	**32PFU**	
SF	**Rift Valley Fever**	MP 12	vaccine Strain	**POS/YES**	**100PFU**	
SF	**Sandfly fever Naples**	Original	human; Naples, Italy, 1944	**POS/YES**	**0.23PFU**	
SF	**Toscana**	IssPhl 3	*Phlebotomus perniciosus* ; Toscana, 1971	**POS/YES**	**0.95PFU**	
SF	**Sandfly fever Sicillian**	Original	human; Sicily, 1943	**POS/YES**	**0.0045PFU**	
SF	Cacao*	VP 437R	*Lutzomyia trapidoi* ; Panama, 1970	**POS/YES**	ND	
SF	Buenaventura	CoAr 3319	*Lutzomyia spp.;* Colombia, *1964*	**POS/YES**	ND	
SF	Frijoles	VP 161A	*Lutzomyia spp.;* Panama, *1969*	**POS/YES**	ND	
SF	Itaituba	bean 213542	*Didelphis masupialis* ; Brazil, 1971	**POS/YES**	ND	
SF	Itaporanga*	Original	sentinal mouse*:* Brazil, *1962*	**POS/YES**	ND	
SF	Gabek Forest*	SudAn 754-61	*Acomys albigena* ; Sudan, 1961	**POS/YES**	ND	
SF	Turuna*	Bear352492	*Lutzomyia spp.* ; Brazil, 1978	**POS/YES**	ND	
SF	Rio Grande*	TBM3-204	*Neotoma micropus* ; Texas, U.S.A., 1973	**POS/YES**	ND	
SF	Nique*	9c	*Lutzomyia panamensis* ; Panama, 1972	**POS/YES**	ND	

[a] Known human pathogens appear in bold.

[b] Abbreviations include CCHF, Crimean Congo hemorrhagic fever group; NSD, Nairobi sheep disease group; SF, Sandfly fever group; CAR, Central African Republic; PFU, plaque forming unit equivalent; ND, not determined.

[c] PFU equivalent determined through plaque titration of virus dilutions in Vero cells.

[d] Garissa and Germiston viruses were not evaluted due to a lack of availability (Garissa) and handling restrictions (Germiston).

[e] *S segment sequence data were determined for this virus prior to this study.

Table 3.2. Arthropod-borne human pathogens that are not amplified by the molecular consensus assay[a]

Family/ *Genus*	***Serogroup***	***Virus***[b]	***Strain***	***source, location and date of isolation***
Bunyaviridae				
Orthobunyavirus	*Nyando*	**Nyando**	MP 401	*Eretmapodites chrysogaster* ; Cameroon, 1966
Orthobunyavirus	*California*	**Guaroa**	CoH 352111	human; Colombia, 1956
Orthobunyavirus	*Group C*	**Apeu**	BeAn 848	*Cebus apella* ; Brazil, 1955
Orthobunyavirus	*Group C*	**Caraparu**	BeAn 3994	*Cebus apella* ; Brazil, 1956
Orthobunyavirus	*Group C*	**Itaqui**	BeAn 1297	sentinal mouse; Brazil, 1959
Orthobunyavirus	*Group C*	**Madrid**	BT 4075	human; Panama, 1961
Orthobunyavirus	*Group C*	**Marituba**	BeAn 15	*Cebus apella* ; Brazil, 1954
Orthobunyavirus	*Group C*	**Nepuyo**	BeAn 10709	sentinal mouse; Brazil, 1959
Orthobunyavirus	*Group C*	**Oriboca**	BeAn 17	sentinel monkey; Brazil, 1954
Orthobunyavirus	*Group C*	**Restan**	TRVL 51144	*Culex spp* .; Trinidad, 1963
Orthobunyavirus	*Guama*	**Catu**	BeH 151	human; Brazil, 1955
Orthobunyavirus	*Guama*	**Guama**	BeAn 277	*Cebus apella* ; Brazil, 1955
Flaviviridae				
Flavivirus	*Dengue virus group*	**Dengue 2**	New Guinea C	human; New Guinea, 1944
Flavivirus	*Yellow fever virus group*	**Yellow Fever**	17D	vaccine strain
Flavivirus	*Spondweni virus group*	**Zika**	YAP-human, 200	RNA transcript derived from PCR amplified DNA
Flavivirus	*Japanese encephalitis group*	**SLE**	TBH-28	human; Florida, U.S.A., 1962
Flavivirus	*Japanese encephalitis group*	**WNV**	NY99	*Phoenicopterus ruber* ; New York, U.S.A., 1999
Flavivirus	*Japanese encephalitis group*	**JE**	SA-14-2-8	vaccine strain
Togaviridae				
Alphavirus	*Venezualan Equine encephailitis Complex*	**VEEV**	TC83	vaccine strain
Alphavirus	*Western Equine encephalitis complex*	***WEEV***	McMillan	human; Ontario, Canada, 1941
Alphavirus	*Semliki Forest complex*	***ONN***	UgMP 30	human; Uganda, 1959
Alphavirus	*Semliki Forest complex*	***CHIKV***	Original	human; Tanganyika, 1953

[a] All viruses were evaluated with both the Orthobunyavirus and Phlebo/Nairovirus primer mixes.

[b] All evaluated viruses have a minimum titer of 10000 PFU/mL in Vero cells.

Recently derived orthobunyavirus isolates of international origins. A panel of recent orthobunyavirus isolates was assembled to evaluate the presented assays ability to detect target RNAs from a diversity of origins (Table 3.3). Isolates from Africa and China were derived from mosquito pools and described and characterized by others prior to their evaluation in our laboratory (Crabtree et al, 2009; Lu et al, 2009).

RNA extraction. Viral RNA was extracted from virus seed using the QIAamp Viral RNA mini kit (Qiagen). For specificity testing, extractions were performed on samples ranging in volume from 70-140 µl. RNA was eluted in a volume equal to the volume of starting sample. For sensitivity analysis, extractions were performed on plaque-titrated virus dilutions of 220µl in volume. RNA was eluted in 110µl of elution buffer (Qiagen) for a 2X concentration of RNA. Eluted RNA was stored at -70ºC until evaluated by RT-PCR amplification.

Table 3.3. Detection and identification of recently derived orthobunyavirus isolates by the molecular consensus assay [a]

Genus/ *Serogroup*	***Virus***	***source, location and date of isolation***	**RT-PCR result / serological *identity confirmed by nucleotide sequencing***
Orthobunyavirus			
Bunyamwera	**Bunyamwera**	*Aedes ochraceus;* Kenya, 2006	**POS/YES**
Bunyamwera	**Bunyamwera**	*Aedes ochraceus;* Kenya, 2006	**POS/YES**
Bunyamwera	**Bunyamwera**	*Aedes ochraceus;* Kenya, 2006	**POS/YES**
Bunyamwera	**Bunyamwera**	*Aedes ochraceus;* Kenya, 2006	**POS/YES**
Bunyamwera	**Bunyamwera**	*Aedes ochraceus;* Kenya, 2006	**POS/YES**
Bwamba	**Pongola**	*Aedes mcintoshi;* Kenya, 2006	**POS/YES**
California	**Tahyna**	mosquito *spp;* China, 2007	**POS/YES**
Flavivirus	**Zika**	human; YAP, 2007	NEG
Alphavirus	**Chikungunya**	human*;* returning traveller to USA, 2006	NEG

[a]Viruses were isolated and serologically characterized at the CDC as part of international collaborations, with the exception of the Tahyna virus isolate from China. Tahyna viral RNA, extracted from Tahyna virus isolated and serologically characterized in China, was submitted to the CDC for confirmatory testing.

Primer design. For amplification and sequencing, oligonucleotide primers (Table 3.4) were designed to target previously determined and newly derived S segment sequences of human pathogens of the *Orthobunyavirus, Phlebovirus,* and *Nairovirus* genera of the family *Bunyaviridae* (Tables 3. 1 & 3.2). S segment sequences were aligned by serogroup and genus, using MegAlign software (DNA*). According to these alignments, primers were designed manually to target conserved regions shared within the overlapping N and NSs ORFs of analyzed orthobunyavirus sequences and the N ORF of analyzed nairo- and phlebovirus sequences through visual inspection (Figure 3.1; Table 3.4).

RT-PCR molecular consensus assay amplification of bunyaviral cDNAs. For the detection and amplification of medically important and serologically related viruses of the family *Bunyaviridae*, the following procedures were followed. RNAs extracted from bunyaviruses of interest and other arthropod borne viruses (Tables 3. 1 & 3.2) were subjected to RT-PCR amplification, using a combination of virus and serogroup specific primers designed to detect viruses of the *Orthobunyavirus, Nairovirus* and *Phlebovirus* genera (Tables 3. 1 & 3.2). RT-PCR assays were performed with 20µl of extracted viral

Table 3.4. Oligonucleotide primers for the amplification and sequencing of the partial S segment of viruses of the family *Bunyaviridae*

Targetted genus/genomic target	primer name	***Orthobunyavirus* forward primer mix**	primer name	***Orthobunyavirus* reverse primer mix**	Approx. amplicon size
Orthobunyavirus N ORF	Cal/Bwa group forward	GCAAATGGATTTGATCCTGATGCAG	Cal/Bwa group reverse	TTGTTCCTGTTTGCTGGAAAATGAT	210
Orthobunyavirus N ORF	Bun group forward	CTGCTAACACCAGCAGTACTTTTGAC	Bun group reverse	TGGAGGGTAAGACCATCGTCAGGAACTG	250
Orthobunyavirus N ORF	Wyeomyia forward	ATGTCTGAAATTGTATTTGATGATATTGG	Wyeomyia reverse	TATTTCGATTCCCCGGAAAGT	230
Orthobunyavirus N ORF	Oropouche forward	GGCCCATGGTTGACCTTACTTT	Oropouche reverse	ACCAAAGGGAAGAAAGTGAAT	300
		***Phlebo/Nairo* forward primer mix**		***Phlebo/Nairo* reverse primer mix**	
Nairovirus S segment N ORF	Nairo forward	TCTCAAAGAAACACGTGCCGC	Nairo reverse	GTCCTTCCTCCACTTGWGRGCAGCCTGCTGGTA	400
Phlebovirus N ORF	Phlebo forward 1	TTTGCTTATCAAGGATTTGATGC			
Phlebovirus N ORF	Phlebo forward 2	TTTGCTTATCAAGGATTTGACC	Phlebo reverse	TCAATCAGTCCAGCAAAGCTGGGATGCATCAT	370

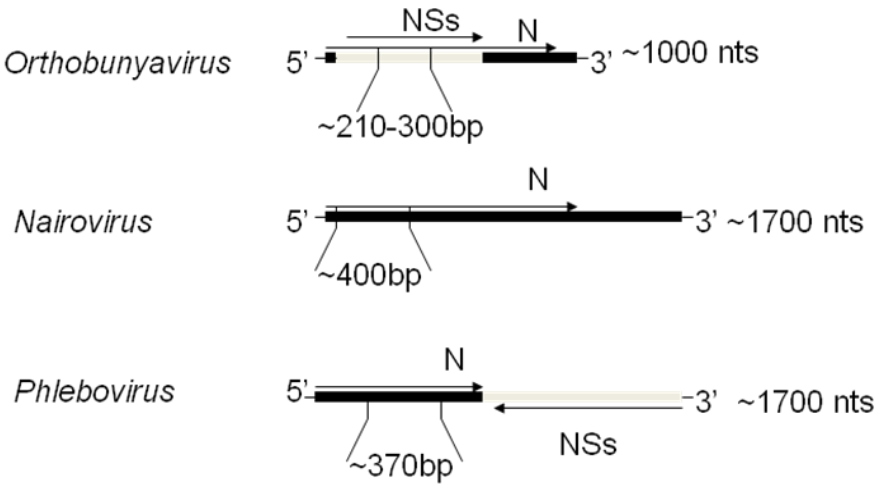

Figure 3.1. Targeted regions of the presented assay shown within S segment vcRNA N and NSs ORF coding strategies of arthropod-borne genera of the family *Bunyaviridae*

RNA, and 50 pmol of each forward and reverse primer listed in mixes for detection of either i) the *Orthobunyavirus* genus or ii) the *Nairovirus and Phlebovirus* genera (Table 3.4) in a 50 µl total reaction volume using Qiagen OneStep RT-PCR mastermix (Qiagen). Reactions were subjected to RT-PCR amplification using the following cycling conditions: one cycle of 50 ºC for 30 minutes, 95 ºC for 15 minutes; followed by 55 cycles of 94 ºC for 30 seconds, 55 ºC for 1 minute and 72 ºC for 2 minutes. Reactions were terminated after a final extension step at 72 ºC for 10 minutes.

Primer selection and assay optimization. Prior to the selection of presented primers (Table 3.4) and assay reaction conditions (see **RT-PCR molecular consensus assay amplification of bunyaviral DNAs)**, numerous alternatives were evaluated (data not shown). The presented oligonucleotide mixes and reaction conditions were selected based on the ability to detect the largest number of bunyaviruses with the highest relative sensitivity.

Consensus assay result determination and DNA purification. Products of RT-PCR amplification were evaluated by electrophoresis in a 2.0% agarose gel in 40mM Tris-Acetate-1mM EDTA (TAE) buffer. Five µl of product were analyzed for each amplified sample. Visualization of DNA bands was achieved through ethidium bromide staining and UV transillumination. A sample was determined to be presumptive bunyavirus-positive through size-determined migration and visualization of DNA bands, ranging from 210-400bp (Figure 3.2; Table 3.4), on the gel. For DNA purification, the remaining 45 µl of a positive sample was directly purified without the need for gel extraction using the QIAgen PCR purification kit (Qiagen). In exception, the presence of multiple bands necessitated gel extraction of the target sized band from the amplification products of Gabek Forest virus, using the QIAgen Gel Extraction kit (Qiagen).

Multiplex nucleotide sequencing of RT-PCR amplified *Orthobunyavirus, Nairovirus* and *Phlebovirus* DNAs for bunyavirus-positive result confirmation and

preliminary species identification. To determine the nucleotide sequence of purified cDNAs, amplified from viral RNAs of the *Orthobunyavirus, Nairovirus and Phlebovirus*

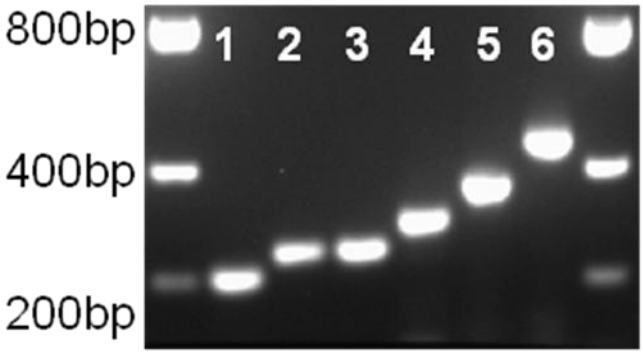

Figure 3.2. Gel image of fragments amplified by the presented consensus assay. Fragments range in size from ~210-400bp. Fragments from the amplification of 1, La Crosse ; 2, Wyeomyia; 3, Bunyamwera; 4, Oropouche; 5, Punta Toro; and 6, Dugbe viruses are represented. Five µl of each DNA was loaded onto a 2% agarose gel containing EtBr stain.

genera, the following steps were taken. DNAs were sequenced with cocktails of forward or reverse primers targeting either, i) the *Orthobunyavirus* genus, or ii) the *Nairovirus* and *Phlebovirus* genera (Table 3.4). Three and two-tenths pmol of each listed primer in the appropriate forward or reverse mix (Table 3.4), along with approximately 50 ng of purified DNA, were added per reaction using the ABI BigDye Terminator V3.1 Ready Reaction Cycle Sequencing mix (Applied Biosystems). Sequencing reactions were purified using the Qiagen DyEx 2.0 spin kit (Qiagen). Nucleotide sequences were determined by running purified sequencing reactions on the ABI 3130 genetic analyzer (Applied Biosystems). Using these methods, partial S segment sequences were generated for all bunyaviruses amplified by the presented assay (Tables 3. 1 & 3.3).

S segment nucleotide sequence analysis of DNAs amplified by the consensus assay and preliminary species identification through Nucleotide BLAST search

mechanism of NCBI GenBank database. Multiple nucleotide sequences generated for each evaluated bunyavirus (Tables 3. 1 & 3.3) were aligned in both the 5' and 3' directions and then edited using SeqMan software (DNA*). The resultant consensus sequence was then queried using the NCBI BLAST function of the Nucleotide Collection database at http://blast.ncbi.nlm.nih.gov, (Altschul et al, 1990). For S segment species identification, queries were optimized for somewhat similar sequences (blastn) using default algorithm parameters. Using these methods, sequences that matched the queried sequence with the highest overall maximum value were considered to provide the closest identity to the S segment species of an evaluated DNA. From these data, preliminary species-level identification of evaluated virus isolates was made.

Results and Discussion

The analytical sensitivity of the bunyavirus molecular consensus assay, as determined through its application to plaque-titrated human pathogens of the family B*unyaviridae,* is reported (Table 3. 1). For the *Orthobunyavirus* genus, the molecular consensus assay detects between 660 PFU/0.1mL (for Wyeomyia virus) to 0.1PFU/0.1mL (for Oropouche virus). For the *Phlebovirus* and *Nairovirus* genera, detection limits range from 100PFU/0.1mL (for Rift Valley Fever virus) to 0.0024PFU/0.1mL (for Chagres virus).

The analytical specificity of the bunyavirus molecular consensus assay, as determined through its application to viruses of the families *Bunyaviridae, Flaviviridae* and *Togaviridae*, is presented (Tables 3. 1 & 3.2). Forty seven viruses, including 29 human pathogens, representing six major serogroups of the family Bunyaviridae were targeted and amplified by the presented assay (Table 3. 1). Demonstrating fidelity for bunyaviral RNA targets, no evaluated viruses of the *Flaviviridae* or *Togaviridae* families were

detected by the consensus assay (Table 3.2). For all amplified bunyaviral DNAs, multiplex sequencing and NCBI BLAST analyses generated results for S segment nucleotide sequence data that confirmed the serologically derived identity, with the exceptions of Ilesha virus K0/2, Maguari virus 75V3429, and Snowshoe hare virus LIEV 13004AKH strains (Table 3. 1). The S segment species amplified and sequenced from the Ilesha virus KO/2, Maguari virus 75V3429, and Snowshoe hare virus LIEV 13004AKH strains shared the highest nucleotide sequence identities with various Batai, Cache Valley and Chatanga virus strains, respectively, according to our analyses (data not shown).

The ability of the bunyavirus consensus assay to detect target viral RNAs extracted from temporally relevant samples was evaluated through its application to a panel of isolates derived from recent, international origins (Table 3.3; Crabtree et al, 2009; Lu et al, 2009). The presented molecular consensus assay amplified and identified Bunyamwera, Pongola and Tahyna viral cDNAS of the Bunyamwera, Bwamba and California serogroups of the genus *Orthobunyavirus* from viruses isolated from Kenya and China, in 2006 and 2007, respectively (Table 3.3). Confirming fidelity for bunyaviral RNAs, other recently derived isolates of the *Alphavirus* and *Flavivirus* genera were not amplified by the presented consensus assay (Table 3.3).

To the best of our knowledge, the presented bunyavirus molecular consensus assay has been shown to detect a larger diversity of viruses than any other molecular assay designed for the detection of arthropod-borne viruses. In addition, the assay has been shown to have sufficient sensitivity for the detection of virus strains of varying titer isolated from a diverse range of host species, geographic and temporal origins (Tables 3.1 & 3.3). Also, when combined with multiplex sequencing and NCBI BLAST analyses, the presented assay was able to confirm serologically derived identities for all but three of 47

amplified viruses (Tables 3.1 & 3.3). Finally, the assay has been shown to have fidelity for bunyaviral RNAs when applied to arthropod-borne viruses of the families *Flaviviridae* and *Togaviridae* (Tables 3.2 & 3.3).

In addition to viruses of the *Phlebovirus* and *Orthobunyavirus* genera that were detected (Tables 3. 1 & 3.3) some evaluated orthobunyaviruses were not amplified by the presented assay (Table 3.2). While we were able to detect Nyando and Guaroa viruses using alternative primer mixes and reaction conditions to those presented here, these alternatives compromised the sensitivity of detection for other medically important agents and, as a result, were abandoned in the final assay design (data not shown). For the Group C and Guama serogroup viruses presented in Table 3.2, no efforts (presented or alternative) generated DNAs for sequencing. These findings are particularly confounding with regard to the Group C agents for which multiple amplification attempts, using oligonucleotide primers designed to target previously published Group C S and M segments (Nunes et al, 2005), failed to generate DNAs for sequencing. Having tried to eliminate other variables that might contribute to disparities in amplification efficiency, it is our belief that target nucleotide sequence heterogeneity is responsible for the inability to detect these agents by molecular methods.

The identification of S segment species through multiplex sequencing and NCBI BLAST analyses is a powerful component of the presented molecular consensus assay. However, we regard the determination of virus species through GenBank derived S segment identity as preliminary for the following reasons. i) It is often impossible to verify the integrity of sequence data for a given virus species in Genbank. As a result, inaccurate data could potentially cause the misidentification of an S segment species using the presented methodologies. However, our collective experience using GenBank

queries for virus identification has indicated a high degree of integrity of these data in instances where we were able to verify virus identity by alternative methods (Campbell et al, 2006; Lanciotti et al, 1999; Lanciotti et al, 2007; Lanciotti et al, 2008). ii) The capacity for genomic segment reassortment amongst viruses of the family *Bunyaviridae* allows for the possibility that a given isolate possesses genomic segments of heterologous ancestry. Initially, we considered our findings of Cache Valley, Batai and Chatanga virus-like S segments, amplified from isolates serologically classified as Maguari, Ilesha and Snowshoe hare virus strains, respectively (Table 3. 1), evidence of possible segment reassortment. However, subsequent analyses performed on multi-segment data generated for these viruses (data not shown) indicate that inadvertent serological misclassification of Cache Valley, Batai and Chatanga viruses as Maguari, Ilesha and Snowshoe hare viruses is responsible for the disparity in results. Serological misclassification is likely the result of a limited diversity of antibodies in identifying laboratories at the time of their characterization. Finally, because of the above described limitations on virus identification by GenBank query, we recommend that the preliminary identification of a virus species through GenBank derived S segment identity always be confirmed by additional molecular and/or serological analyses.

The molecular consensus assay is generally more sensitive for the detection of viruses of the *Nairovirus* and *Phlebovirus* genera than for the detection of viruses of the *Orthobunyavirus* genus (Table 3. 1). We have considered several potential explanations for this difference in senstitivities. i) From our collective laboratory experience developing and applying assays for the detection of arthropod-borne viruses, it has been noted that, in general, increasing the number of different primers in a reaction causes a reduction in overall sensitivity of detection. To accommodate a greater diversity of

nucleotide sequence targets, the *Orthobunyavirus* mix contains more primers than the *Nairovirus*/*Phlebovirus* mix does, possibly contributing to the noted differences in sensitivities of detection (Table 3. 1). ii) Oligonucleotides designed for the detection of orthobunyaviruses bind target nucleic acid less efficiently than those designed for the detection of nairo- and phleboviruses. In fact, to incorporate a broader diversity of targets, primers designed for the detection of viruses of the *Orthobunyavirus* genus are generally less specific for target sequence than are those for the detection of viruses of the *Nairovirus* and *Phlebovirus* genera. Therefore, the O*rthobunyavirus* primers (Table 3.4) likely prime RT-PCR amplification less efficiently, resulting in reduced sensitivity. iii) There could be a difference in the ratio of RNA copies/infectious viral particles in the Vero cell plaque assay for these viruses. If the evaluated phlebo- and nairoviruses have a higher ratio of RNA copies/PFU equivalents in the Vero cell culture system than that of the evaluated orthobunyaviruses, it would appear that the sensitivity of detection for these agents is greater. The application of this assay to standards of known RNA copy number is supported by this conjecture. However, because the PFU is the most relevant unit for quantitating infectious virus, we consider experiments designed to determine assay sensitivity, through its application to standards of known copy number, beyond the necessary scope of this research.

The data presented here (Tables 3.1 & 3.3) demonstrate the utility of the molecular consensus assay for the detection of virus isolates of the family *Bunyaviridae.* Most promising to us is the ability of the assay to correctly identify recently derived human pathogens of the genus *Orthobunyavirus,* isolated from mosquito pools as part of arbovirus surveillance programs in Kenya and China (Crabtree et al, 2009; Lu et al, 2009; Table 3.3). Unfortunately, we do not have a comparable panel of recent isolates for the

characterization of the *Nairo-/Phlebovirus* component of this assay. However, it is believed that the extremely high senstivities reported for the detection of historically obtained viruses of these genera (Table 3. 1) are conducive for the detection of newly emergent nairo- and phleboviruses as well. We recommend the incorporation of the presented molecular consensus assay in an algorithm that includes additional molecular and serological methodologies for the comprehensive identification of medically important, and serologically related, virus isolates of the family *Bunyaviridae.* Additionally, it is our intention that the presented assay might be combined with other previously published (Honig et al, 2004) or newly derived assays to target multiple segments of the bunyavirus genome. However, research to determine the feasibility of an assay with a designed capability for the detection of reassortant bunyaviruses was not performed in this study. Finally, because this assay has been designed to target conserved regions shared between genetically disparate agents, we believe that additional bunyavirus species to those evaluated here will be amplified and detected by the presented assay. However, the paucity of genetic information available for many of these agents and extraordinary diversity amongst viruses of the family *Bunyaviridae* prevents absolute certainty that the assay will amplify uncharacterized viruses.

CHAPTER 4

THE APPLICATION OF NEWLY DEVELOPED METHODS FOR THE DETECTION OF BUNYAVIRUSES OF NOVEL GEOGRAPHIC AND HOST ORIGINS

Abstract

The primary applied public health goal of the research described within this dissertation has been to enhance the capability for the detection and identification of arthropod-borne viruses of the family *Bunyaviridae.* To that end, the determination of primary nucleotide sequence data for viruses of previously undescribed nucelotide sequences and the resultant development of a broadly reactive consensus assay have been described (Chapters 2&3). Here, the application of the understanding and methods developed in Chapters 2 & 3 for the detection of bunyaviruses of diverse classification, host and geographic origins is presented. The subsequent discoveries of Tahyna virus in China in 2006, La Crosse virus in *Aedes albopictus* mosquitoes in Dallas County, Texas in 2009 and Toscana viral RNA within the CSF of an American traveler who returned from a trip to Sicily in 2009 are described.

Introduction

At the Centers for Disease Control and Prevention (CDC), Division of Vector-Borne Diseases, Arthropod-Borne Diseases branch, Diagnostic and Reference laboratory we have a specific mission to identify arthropod-borne viruses in environmental and clinical samples that are submitted to us by referring laboratories from around the world. In support of this mission, we have applied newly developed molecular methods (Chapters 2 &3; Lambert & Lanciotti, 2008; Lambert & Lanciotti, 2009) for the detection of bunyaviruses of diverse origins. Here, the detection and characterization of Tahyna, La Crosse and Toscana viral RNAs in samples submitted to us from China, Texas, USA and an American traveler that returned from Sicily to Washington State, USA, respectively, are described (Kay et al., 2010; Lambert et al., 2010; Lu et al., 2009). These discoveries underscore the immediate and extraordinary utility of the methods described within this dissertation (Chapters 2 &3; Lambert & Lanciotti, 2008; Lambert & Lanciotti, 2009) and are a reflection of their current and future value to international diagnostic and epidemiological investigations.

Tahyna virus in China, 2006

First isolated from a pool of *Aedes caspius* mosquitoes in Czechoslovakia in 1958 (Bardos et al., 1959), Tahyna virus has been identified as a cause of human illness throughout Europe and the former Soviet Union since that time (Gould et al., 2006; Bulychev et al., 1985; Lvov et al., 1977; Hubalek et al, 2010). A member of the California serogroup of the genus *Orthobunyavirus,* Tahyna virus is vectored by *Aedes spp* mosquitoes with small mammals, including rabbits and domestic animals, serving as amplifying hosts. Humans are dead-end hosts. Upon infection with Tahyna virus, human

clinical manifestations can include influenza–like respiratory symptoms with CNS complications occurring in rare cases. Historically, there has been no known occurrence of Tahyna virus or Tahyna virus-related illness in China.

As part of arthropod-borne surveillance efforts undertaken in the People's Republic of China in the summer of 2006, mosquito pools were collected in light traps in the Kashi area of the Xinjiang Uygur Autonomous Region of China near the Kyrgyzstan boarder (Lu et al., 2009; Figure 4.1). Collected mosquitoes were identified and separated by genus/species/collection site in pools containing up to 100 specimens at the Chinese Center for Disease Control and Prevention in Beijing, China (Lu et al., 2009).

Figure 4.1 A) Map of China indicating the Xinjiang Uygur Autonomous Region, which appears shaded. B) Kashi region (star) of the Xinjiang Uygur Autonomous Region where the mosquitoes that comprise the pool that yielded the XJ0625 isolate that was determined to be Tahyna virus were collected. Figure adapted from Lu et al., 2009.

essential medium, clarified by centrifugation and the resultant supernatants were inoculated onto BHK-21 and Vero cells in 6 well plates (Lu et al., 2009). Inoculated plates were then incubated at 37°C and reviewed for cytopathic effect daily (Lu et al., 2009). Both BHK-21 and Vero cells that were inoculated with the supernatant, generated as described above, from a single pool containing 100 *Culex spp* mosquitoes that were both unfed and blood-fed demonstrated significant cytopathic effect within two days of inoculation (Lu et al., 2009). The resultant isolates were preliminarily characterized as California serogroup viruses by both molecular and serological methods in China. To further identify the BHK-21 isolate at the molecular level, RNA extracted from this isolate, XJ0625 was sent to our laboratory at the CDC in Fort Collins, CO in April of 2008.

Upon receipt at the CDC, molecular methods previously described for the amplification and sequencing of full-length S segment cDNAs from viruses of the genus *Orthobunyavirus* (Chapter 2; Lambert & Lanciotti, 2008) were applied to the XJ0625 RNA sample. The resultant full-length S segment nucleotide sequence was identified as Tahyna viral in origin by NCBI BLAST analyses (Altschul et al., 1990) performed as previously described (Chapter 3; Lambert & Lanciotti, 2009). Subsequent analyses, conducted at the submitting laboratory in China, confirmed the identities of the M and L segments of XJ0625 as Tahyna viral in origin as well (unpublished data). This discovery represents the first known occurrence of Tahyna virus in the People's Republic of China; however, the public health impact of this virus in China remains unknown.

La Crosse virus in Aedes albopictus mosquitoes; Dallas County Texas, USA 2010

La Crosse virus (LACV) is the most important cause of arthropod-borne, pediatric encephalitis in North America. A member of the California serogroup of the genus *Orthobunyavirus,* LACV is associated with neuroinvasive illness that can be fatal and occurs within the geographic range of the principal vector, *Aedes triseriatus.* A native tree-hole breeding mosquito, *Aedes triseriatus* is widely distributed throughout wooded regions east of and bordering the Mississipi River within the United States. Historically, the majority of LAC encephalitis cases have occurred in upper midwestern states including Wisconsin, Illinois, Minnesota, Indiana and Ohio (Figure 4.2). In recent years, LAC encephalitis activity has increased above endemic levels in regions of the southeastern United States, including West Virginia, North Carolina and Tennessee (Haddow & Odoi, 2009; Gerhardt et al., 2001; Figure 4.2). In addition, recent cases of La Crosse encephalitis have been reported as far south as Louisiana, Alabama, Georgia and Florida (Figure 4.2).

Aedes albopictus is an invasive mosquito species that was first discovered in Houston, Texas in 1985 (Centers for Disease Control, 1986); having apparently arrived to the United States in a shipment of used tires from Asia (Moore & Mitchell, 1997).
An opportunistic container-breeder, its vector competence for many arthropod-borne viruses, including LACV, and a catholic feeding habit have made the invasion of *Aedes albopictus* disconcerting to researchers who have warned of the potential for an increased incidence of vector-borne disease as a result (Grimstad et al., 1989; Weaver & Reisen, 2010). Since 1985, the geographic distribution of *Aedes albopictus* has grown

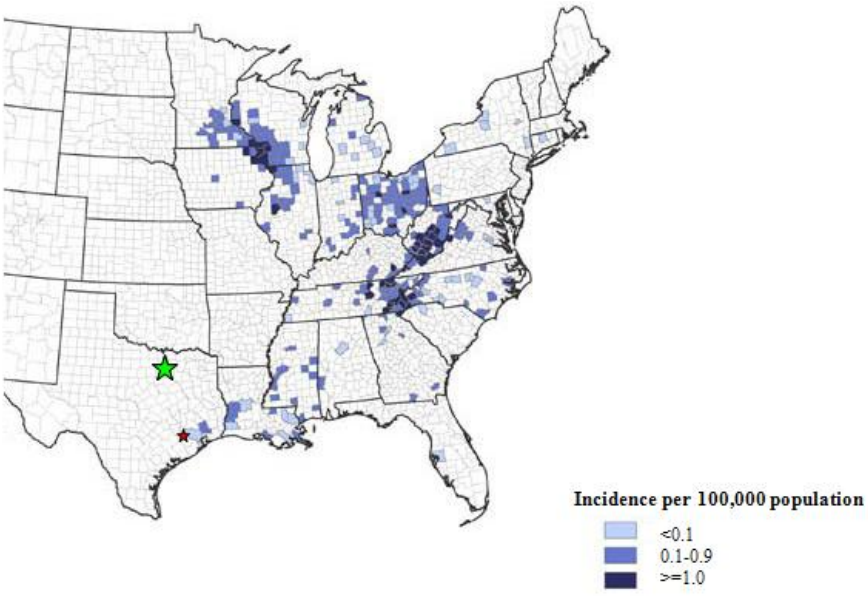

Figure 4.2 Geographic distribution of La Crosse virus in accordance with the habitat range of *Aedes triseriatus* in the United States as inferred from the California serogroup virus neuroinvasive disease average annual incidence by County; 1996-2008. Incidence rates shown in shades of blue. Dallas County and Fort Bend County locations of the 2009 LACV isolations from pools containing *Aedes albopictus* and *Aedes triseriatus* mosquitoes are indicated by green and red stars, respectively. Presented data and figure adapted from the Centers for Disease Control and Prevention website at www.cdc.gov and Lambert et al., 2010.

to include the majority of the southeastern United States. The concurrent increase in LAC encephalitis activity in the southeastern United States (Haddow & Odoi, 2007) has engendered speculation on the possible transmission of LACV by *Aedes albopictus* as an accessory mechanism to the historically understood transmission by *Aedes triseriatus* (Gerhardt et al., 2001). In fact, LACV has been isolated from *Aedes albopictus* mosquitoes in Tennessee and North Carolina in 1999 and 2000, respectively, during a period of greatly increased LACV activity in those areas (Gerhardt et al., 2001). However, the role of *Aedes albopictus* in LACV transmission in nature remains unknown.

The isolation of La Crosse virus from a pool of three *Aedes albopictus* mosquitoes that were collected in Dallas County, Texas on August 13, 2009 is reported (Figure 4.2). To the best of our knowledge, this is the first report of an endemic arthropod-borne virus occurring outside its known geographic range within an invasive vector species in the United States. Additionally, this is only the second isolation of LACV within the state of Texas; the first isolate having been derived from a pool of *Aedes infirmatus* mosquitoes that were collected in the City of Houston over thirty eight years ago, in 1972 (Klimas et al., 1981). Of interest, subsequent to the identification of LACV in the Dallas pool, an additional isolation of LACV was made from a mixed pool of 29 *Aedes albopictus* and two *Aedes triseriatus* mosquitoes that were collected in Fort Bend County, Texas in October of 2009 (Figure 4.2). The Fort Bend County location is relatively near the site of collection of the 1972 Texas LACV-positive pool and the known geographic distribution of LACV activity in southeastern Texas and Louisiana (Figure 4.2). Taken together, our results represent an unprecedented number of findings of LACV within the state of Texas.

As part of arthropod-borne virus surveillance efforts, the City of Dallas Vector Control division collected 65 mosquitoes in a gravid trap at the edge of a wooded area near a residential district in Dallas County, Texas on August 13, 2009. Upon receipt at the Texas State Department of Health Services, the mosquitoes were sorted and identified by sex. Identified female mosquitoes were then grouped into three pools by species: pool ID# AR6318 consisting of 50 *Culex quinquefasciatus* mosquitoes, pool ID#AR6319 consisting of three *Aedes albopictus* mosquitoes and pool ID#6320 consisting of one *Aedes triseriatus* mosquito. None of the mosquitoes was engorged with blood.

Generated pools were macerated in 1mL of BA-1 medium followed by centrifugation and inoculation of the resultant supernatant onto BHK and Vero cells. Inoculated cells were incubated at 37º C and were reviewed for cytopathic effect over the course of 10 days. At day five post-inoculation, Vero cells that were inoculated with the supernatant derived from pool ID#AR6319 *(Aedes albopictus)* demonstrated significant cytopathic effect, representing a preliminary virus isolation-positive result. Infected cells were then subjected to IFA utilizing antibodies directed against various arthropod-borne viruses. From these analyses the isolate derived from pool ID#AR6319 *(Aedes albopictus)* was determined to be a California serogroup virus. Of additional note, Pool ID#6318 (*Culex quinquefasciatus*) was identified as positive for West Nile virus and Pool ID#6320 (*Aedes triseriatus*) was identified as negative for virus presence by the above described methods.

In an attempt to further describe the California serogroup virus identified in pool ID# AR6319 *(Aedes albopictus)*, the pool and the derived Vero cell isolate were sent to our laboratory at the CDC in Fort Collins, CO for further testing. Upon receipt at the CDC, an RT-PCR based assay that represents an expansion of the previously described assay (Chapter 3; Lambert & Lanciotti, 2009) was used to amplify cDNAs from all three segments of the *Orthobunyavirus* genome using consensus oligonucleotide primers designed according to methods presented in Chapters 2&3 (Table 4.1) and conditions and methods previously described (Chapter 3; Lambert & Lanciotti, 2009). Generated cDNAs were then subjected to nucleotide sequencing and NCBI BLAST analyses as previously described (Chapter 3; Lambert & Lanciotti, 2009), the results of which indicated that both the pool and the isolate were positive for LACV S, M and L segment RNAs. Subsequently, an additional pool (ID#AR8973) of 29 *Aedes albopictus* and two *Aedes triseriatus*

Table 4.1. Orthobunyavirus consensus oligonucleotide primers used for the amplification and sequencing of La Crosse viral partial S, M, and L segment cDNAs in this study*

Targeted genomic region(s)	Forward primer Name	Forward primer Sequence	Reverse primer Name	Reverse primer Sequence	Approx. amplicon size (bp)
S segment nucleocapsid ORF	Cal S forward	GCAAATGGATTTGATCCTGATGCAG	Cal S reverse	TTGTTCCTGTTTGCTGGAAAATGAT	210
M segment 5' terminus/ glycoprotein ORF	Ortho M 5' terminus	AGTAGTGTACTACC	Ortho M ORF reverse	TTRAARCADGCATGGAA	410
L segment 5' terminus/ polymerase ORF	Ortho L 5' terminus	AGTAGTGTACTCCTA	Ortho L ORF reverse	AATTCYTCATCATCA	550

*Oligonucleotide primers designed against conserved regions of the orthobunyavirus genome. S segment primers and reaction conditions appear in Chapter 3 and Lambert & Lanciotti, 2009. M and L segment primers also appear in Chapter 2.

mosquitoes that were collected in Fort Bend County, Texas on October 5, 2009, was identified as positive for LACV S, M and L segment RNAs using the same processing and characterization methods described above.

Following these analyses, full length S, M and L segment genomic sequences (GenBank accession nos. GU591164–9; Appendix 2) were determined for LACV RNAs that were extracted from both LACV-positive pools using oligonucleotide primers that were designed based on the previously published LACV prototype genome (human 1960, GenBank accession nos. EF485030 (S), EF485031 (M) and EF485032 (L)) and methods formerly described (Chapter 2; Lambert & Lanciotti, 2008). Partial M segment sequences were then aligned by CLUSTAL W (Thompson et al., 1994) and both maximum parsimony (MP) and neighbor-joining (NJ) phylogenetic analyses were conducted using 2000 bootstrap replicates with MEGA version 4 software (Satou & Nei, 1987; Eck & Dayhoff, 1966; Felsenstein, 1985; Tamura et al, 2007). Neighbor-joining trees were generated utilizing the Maximum Composite Liklihood method under the Tamura-Nei model of evolution (Tamura & Nei, 1993; Tamura et al., 2004). Trees generated by MP and NJ methods possess highly similar topologies and bootstrap values therefore a NJ tree was chosen for presentation here (Figure 4.3). The presented phylogeny of partial LAC viral M segment sequences (Figure 4.3) indicates that the LACVs present in the Texas, 2009 pools are highly related to LACVs isolated from Alabama, Georgia and

New York of the previously described lineage 2 (Armstrong & Andreadis, 2006) and genotype C (Klimas et al., 1981) designations. These findings suggest a likely southeastern ancestry for the Texas, 2009 LACV isolates.

The presence of LACV in *Aedes albopictus* mosquitoes in Dallas County, Texas in the late summer of 2009 represents the first documented occurrence of the expansion of the known geographic range of an endemic virus potentially facilitated by an invasive vector species within the United States. The coincident presence of LACV in the Fort Bend County area in October of 2009 should be of unique concern to public health officials and practitioners within the state of Texas. Of additional interest, San Angelo virus, which is serologically related to LACV, is known to occur in Texas and has been shown to replicate in and be transovarially transmitted by *Aedes albopictus* mosquitoes (Tesh & Shroyer, 1980), although this virus has no known association with human disease. Cocirculation enables possible reassortment of genomic segments between LACV and San Angelo virus, a phenomenon that has been described for viruses of the California serogroup within *Aedes albopictus* mosquitoes (Cheng et al., 1999) with unknown public health outcomes.

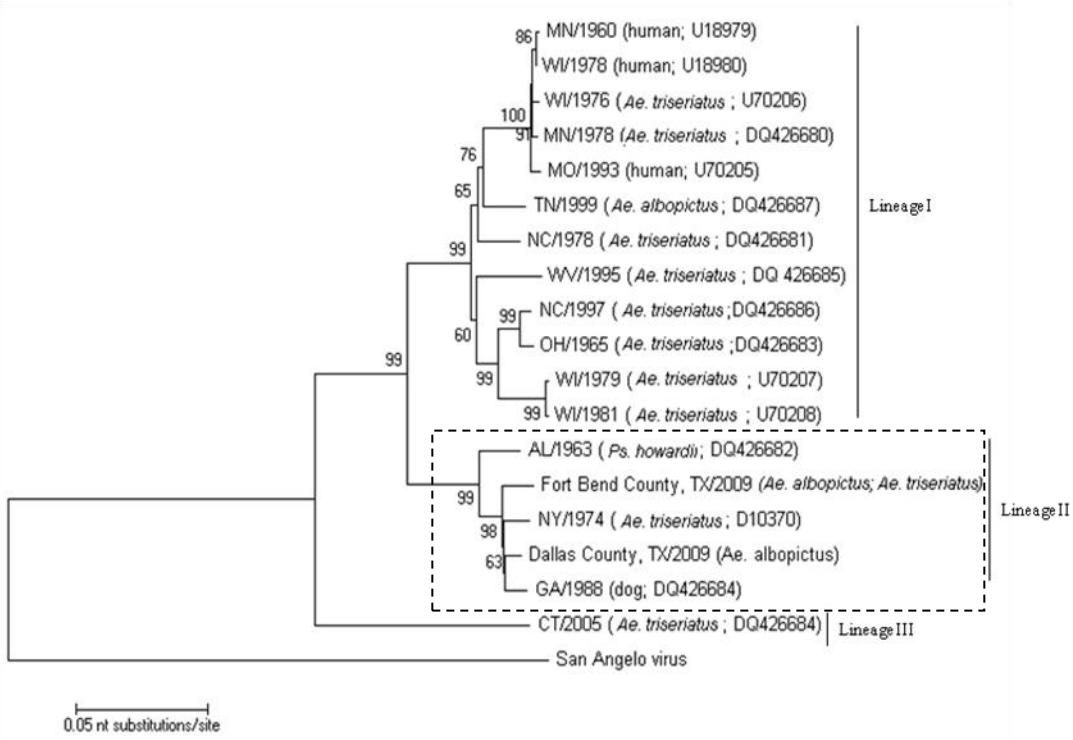

Figure 4.3 Phylogeny of La Crosse virus M segment sequences of diverse origins. According to a limited availability of full length sequences in GenBank, 1663 nts of the M segment glycoprotein gene ORF are compared. Isolate source and GenBank accesssion #s appear after the isolate designation for each taxon. Of interest, the 2009 TX isolates group with extreme support with lineage II viruses of the extreme south and New York suggesting a likely southern origin of the TX LACVs.

Toscana viral RNA in the CSF of an American traveler returning from Sicily in 2009

Toscana virus (TOSV) was first isolated from the sandfly vector *Phlebotomus perniciosus* in central Italy in 1971 (Verani et al., 1982). A member of the genus *Phlebovirus* of the family *Bunyaviridae*, TOSV has become recognized as the leading cause of acute meningitis in central Italy during the summertime (Figure 4.4; Charrel et al., 2005). The most commonly recognized clinical manifestation of TOSV infection is

aseptic meningitis; asymptomatic infections, meningoencephalitis, and encephalitis can also occur (Charrel et al., 2005). Other Mediterranean countries including France, Spain,

Figure 4.4 Region of Italy (encircled) where the predominant number of TOSV human cases have been historically known to occur. Of note, Sicily is remotely located at the southernmost tip of the country. Map is adapted from Big-Italy-Map.co.uk

Of interest, phylogenetic analyses indicate that two geographically distinct genotypes, the Italian and Spanish lineages of TOSV, circulate in the Mediterranean and have been shown to cocirculate in Southeastern France (Sanches-seco et al., 2003). Within Italy, TOSV infections have been known to occur as far south as Naples (Di Nicoulo et al, 2005), and TOSV has been associated with cases of meningitis and encephalitis on the islands of Sardinia and Elba (Venturi et al., 2007; Sonderegger et al., 2009). While TOSV has been identified within sandflies in Sicily (Venturi et al., 2007), to the best of our knowledge, human infection in this southernmost region of Italy has not yet been described. Here, we report a TOSV infection of an American tourist in Sicily in the fall of 2009.

A 65-year-old American physician traveled to Sicily for three weeks and returned to the United States on October 12, 2009. On October 14, he woke with a headache which became progressively worse and he noticed difficulty word-finding (Kay et al., 2010). He was admitted to the hospital that evening (Kay et al., 2010). Upon questioning, it was determined that the patient remained in Sicily during the entire three weeks of his visit, other than changing planes in Milan (Kay et al., 2010). While in Sicily, he explored rural areas and spent time outdoors in the evenings, where he sustained both mosquito and what he thought were flea bites (Kay et al., 2010). According to the patient's clinical presentation, an arboviral encephalatis was considered as the possible etiologic agent and testing of acute and convalescent sera was conducted for evidence of infection with domestic arboviruses at the Washington State Department of Health Laboratory (Kay et al., 2010). Analyses for West Nile virus and St. Louis encephalitis virus immunoglobulin (Ig) M antibodies were subsequently negative (Kay et al., 2010). In light of the patient's recent international travel, other etiologies of viral meningitis were considered. Accordingly, serum and CSF were sent for additional testing at the Centers for Disease Control and Prevention (CDC), Division of Vector-Borne Infectious Diseases, in Fort Collins, Colorado. At the CDC, a preliminary diagnosis of TOSV infection was made by the detection of TOSV RNA in a CSF sample that was collected on day 1 of illness, utilizing the previously described RT-PCR based assay (Chapter 3; Lambert & Lanciotti, 2009). Serological confirmation was provided through the demonstration of a more than four-fold rise in TOSV neutralizing antibodies observed between paired sera collected on days 1 (titer <1:10) and 21 (titer 1:320) of illness by plaque reduction neutralization test analyses (Kay et al., 2010). No similar rise in neutralizing antibodies to serologically related phleboviruses (Sandfly fever Naples and Sandfly fever Sicilian viruses) was

detected by these methods (Kay et al., 2010). The patient had a complete neurological recovery and was able to return to work (Kay et al., 2010).

In light of the novel location of infection, advanced molecular analyses of the S and M segments of TOSV RNA isolated from the infected traveler's CSF were performed at the CDC. These analyses utilized previously published phlebovirus consensus primers that target the S segment of the TOSV genome (Chapter 3; Lambert & Lanciotti, 2009) as well as primers newly designed to target the polyprotein ORF of the M segment of the TOSV genome: TOSV M 851F, 5'ACCAAATACAACCATAGCCCC3' (forward primer) and TOSVM 1327c, 5'ATACAATTCCCACAGTCGTTAG3' (reverse primer). RT-PCR amplification and nucleotide sequencing using the above described primers and methods described previously (Chapter 3; Lambert & Lanciotti, 2009) generated two nucleotide sequences of 332 (S segment) and 424 (M segment) nucleotides in length, Genbank accession nos. GU799570 and GU799571, respectively. Maximum parsimony (MP) and neighbor-joining (NJ) phylogenetic analyses of the newly determined S and M sequences, along with previously determined S and M segment sequences of Mediterranean TOSV isolates of diverse origin were performed using MEGA version 4 software (Tamura et al., 2007). Generated trees share similar topologies and bootstrap values for major Italian and Spanish groupings, therefore a NJ tree generated from M segment sequences of interest was selected for presentation here (Figure 4.5).

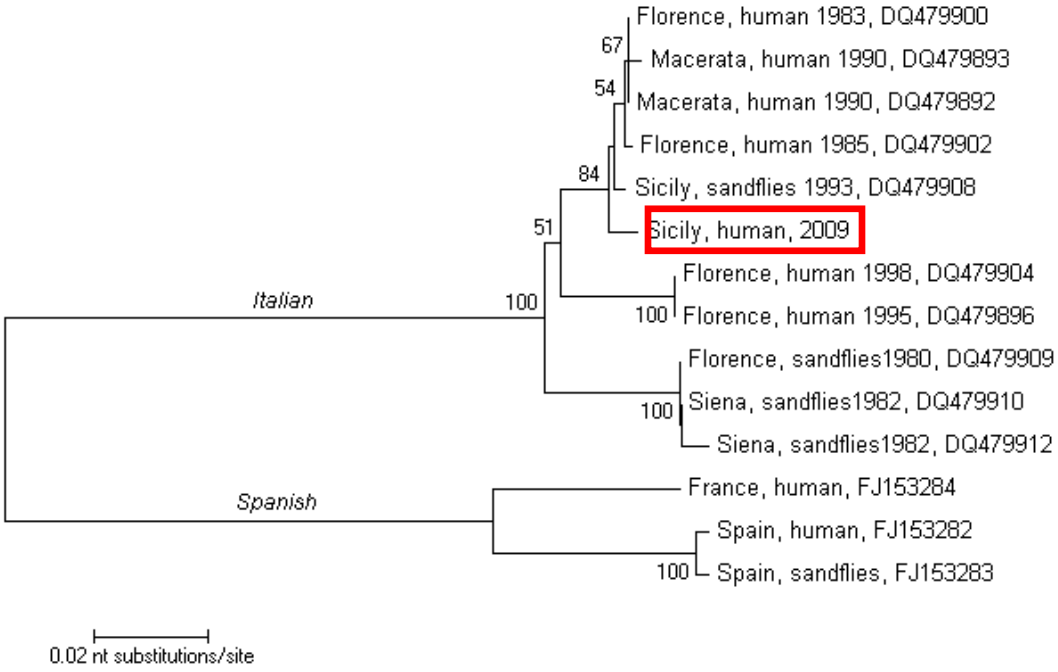

Figure 4.5 Phylogeny of Toscana viruses of diverse origin. A neighbor-joining tree generated from a comparision of 424 nts of the M segment polyprotein gene ORF is shown here. GenBank accesssion nos. appear after the location and source of isolation for each taxon.

According to these analyses, the TOSV RNA identified in the returning traveler described in this study is of the Italian lineage (Figure 4.5). Of interest, the TOSV M segment sequence generated from this patient groups with strong support along with an M segment sequence generated previously from a strain of TOSV that was isolated from sand flies in Palermo, Sicily in 1993 (Figure 4.5). These data indicate that the infecting strain is likely representative of strains that have circulated in Sicily for a number of years.

This case represents the third report of meningitis or meningoencephalitis caused by TOSV infection in an American traveler to Italy (Calisher et al., 1987; Venturi et al., 2007), and the first documented case acquired in Sicily. As is evidenced by this and other

recent reports of TOSV infections in the Mediterranean islands surrounding Italy (Di Nicoulo et al., 2005; Sonderegger et al., 2008), the geographic range of human TOSV infections is greater than previously known. Of additional interest, cases of TOSV infection in travelers returning from endemic regions have contributed to the occurrence and documentation of TOSV-related illness on a global scale (Dobler et al., 1997; Eitrem et al., 1990). Still, many cases of TOSV infection likely go undetected. Studies of seroprevalence in Italy, Spain, Cyprus, and Greece indicate that the proportion of healthy individuals with TOSV IgG antibodies ranges from 20-60% (Charrel et al., 2005); suggesting that the impact of TOSV on public health in the Mediterranean is likely underestimated.

CHAPTER 5

THE MOLECULAR CHARACTERIZATION OF MEDICALLY IMPORTANT VIRUSES OF THE GENUS *ORTHOBUNYAVIRUS*

Abstract

The S segment sequences of Fort Sherman, Shokwe and Xingu viruses of the Bunyamwera serogroup, as well as those of Bwamba and Pongola viruses of the Bwamba serogroup generated in support of Specific Aim 1 (Chapter 2), are analyzed here. S segment sequences of Bwamba and Pongola viruses represent the first nucleotide sequences characterized for viruses of the Bwamba serogroup. The following molecular and phylogenetic analyses of these, and other selected, viruses of the genus *Orthobunyavirus* reveal similarity between the Afrotropical Bwamba and the predominantly Nearctic and Palearctic California serogroups of the genus *Orthobunyavirus.*

Introduction

Orthobunyaviruses are distributed throughout both hemispheres with various small mammals and mosquito species generally serving as amplifying hosts and principal vectors of transmission, respectively (Barrett & Shope, 2005; Calisher & Karabatsos, 1988; Elliott et al., 2000). The majority of human pathogens within the genus *Orthobunyavirus* are classified within three serogroups: i) the predominantly Nearctic and Palearctic California serogroup, ii) the predominantly Neotropical Group C viruses and iii) the predominantly Afro- and Neotropical Bunyamwera serogroup (Barrett & Shope, 2005; Calisher & Karabatsos, 1988; Elliott et al., 2000; Shope & Causey, 1962). To date, the molecular understanding of orthobunyaviruses is limited by a lack of available nucleotide sequence data.

To enhance the molecular understanding of these viruses, we determined the full length S segment RNA sequences of human pathogens of the genus *Orthobunyavirus* (Chapter 2). S segment sequences of Fort Sherman, Shokwe and Xingu viruses of the Bunyamwera serogroup, as well as those of Bwamba and Pongola viruses, the only known members of the relatively small, exclusively African, Bwamba serogroup are described here. Human infections with these viruses have generally been associated with febrile illnesses that can present in a nondescript fashion (associated with Fort Sherman, Shokwe and Xingu viruses), or with arthritis (associated with Pongola virus) or rash (associated with Bwamba virus). Subsequent characterizations and phylogenetic analyses of newly determined sequences in comparison to previously established S segment sequences of selected orthobunyaviruses indicate that the Afrotropical Bwamba serogroup viruses are more closely related to Nearctic and Palearctic viruses of the California serogroup than to other Afrotropical viruses of the genus *Orthobunyavirus*.

Materials and Methods

Characterization of S segment sequences. S segment sequences of Fort Sherman 86MSP18 strain, isolated from a human in Panama in 1985; Shokwe SAAr 4042 strain isolated from *Aedes cuminsii* in Ndumu, South Africa in 1962; Xingu BeH 388464 strain, provided to the WHO reference collection in 1987 and isolated from a human in Brazil on unknown date; Bwamba M459 strain, isolated from a human in Uganda in 1937; and Pongola SAAr1 strain isolated from *Aedes (N.) circumluteolus* in the Pongola river area of South Africa in 1955, generated and confirmed by methods described in Chapter 2, were further characterized by the following methods. Coding and non-coding region sequence lengths and overall A+U content were determined for newly derived and selected previously published sequences of viruses of the genus *Orthobunyavirus* using EditSeq software (DNA*; Table 5.1). Alignments were made using the Megalign software (DNA*) Clustal W function (Thompson et al, 1994) and percent identities were calculated through the use of the MEGA version 4 software pairwise distance calculation function using a "number of differences only" model for both nucleotide and amino acid analyses (Tamura, Dudley, Nei, & Kumar, 2007; Table 5.2).

Table 5.1. Characteristics of newly derived Bwamba and Bunyamwera serogroup S segments shown with those of previously derived full length S segments of representative human pathogens of the California, Bunyamwera, Group C and Simbu serogroups*

Serogroup/virus	Primary biogeographic region	Nucleotide Sequence Length total	5'NCR‖	3'NCR	A+U content mol%	Coding Sequence Length N ORF‡	NSsORF‡	N protein§	NSsprotein§
California									
La Crosse (EF485033)	Nearctic	984	81	195	59.04	**708**†	**279**	**235**	**92**
California Encephalitis (CEU12797)	Nearctic	978	78	192	58.38	**708**	**294**	**235**	**97**
Snowshoe hare (J02390)	Nearctic	982	79	195	56.31	**708**	**279**	**235**	**92**
Inkoo (Z68496)	Palearctic	986	72	206	60.24	**708**	**279**	**235**	**92**
Tahyna (Z68497)	Palearctic	977	78	191	57.73	**708**	**294**	**235**	**97**
Bwamba									
Bwamba (EU564827)	Afrotropical	1096	81	310	64.78	**705**	**279**	**234**	**92**
Pongola (EU564828)	Afrotropical	921	82	131	60.8	**708**	**279**	**235**	**92**
Bunyamwera									
Bunyamwera(NC_001927)	Afrotropical	961	85	174	58.17	702	306	233	101
Shokwe (EU564831)	Afrotropical	958	85	172	58.14	702	306	233	101
Nyando (AM709781)	Afrotropical	959	79	178	59.54	702	306	233	101
Ilesha (AM709780)	Afrotropical	961	85	174	58.79	702	330	233	109
Germiston (BLCSA)	Afrotropical	980	77	191	57.65	702	330	233	109
Cache Valley (X73465)	Nearctic/ Neotropical	950	74	174	59.79	702	306	233	101
Fort Sherman (EU564829)	Nearctic/ Neotropical	948	74	172	59.39	702	306	233	101
Xingu (EU564830)	Neotropical	942	56	184	60.3	702	306	233	101
Group C									
Apeu (DQ188952)	Neotropical	916	76	135	57.5	705	297	234	98
Simbu									
Oropouche(AY237111)	Neotropical	754	45	14	53.5	696	276	231	91

*GenBank accession numbers appear next to virus names. Apeu and Oropouche viruses are included as Group C and Simbu serogroup references, respectively.

†Emboldened values represent highlighted comparisons between the Bwamba and California serogroups.

‡Values are lengths in nt.

§ Values are lenths in aa.

‖NCR, non-coding region

Table 5.2. N gene nucleotide and amino acid sequence identities among human pathogenic California, Bwamba, Bunyamwera,Group C and Simbu serogroup viruses§

Nucleotide or amino acid identity (%)*

		California serogroup					Bwamba serogroup		Bunyamwera serogroup								Group C	Simbu
	Virus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	La Crosse (EF485033)		83.5	88.8	80.1	83.8	**70.6**	**71.0**	53.5	54.5	54.2	50.7	57.1	55.2	55.2	55.7	56.5	52.3
2	California Encephalitis (CEU12797)	85.6		82.5	79.4	83.9	**70.3**	**70.1**	55.2	55.1	56.5	53.3	58.7	54.9	54.5	55.4	59.0	55.9
3	Snowshoe hare (J02390)	90.8	89.1		78.3	83.3	**70.1**	**70.0**	52.5	54.1	52.9	50.4	55.8	54.5	55.4	54.8	56.8	52.6
4	Inkoo (Z68496)	83.0	85.2	82.1		80.3	**68.7**	**70.7**	54.9	56.1	54.9	52.0	55.1	54.5	54.2	57.8	56.7	54.5
5	Tahyna (Z68497)	85.2	89.5	88.2	83.0		**70.9**	**69.9**	54.9	54.6	56.1	53.2	55.5	54.9	54.8	56.7	58.1	53.6
6	Bwamba (EU564827)	**66.4**	**68.1**	**68.6**	**65.1**	**65.**†		*79.9*	56.7	57.2	57.2	54.5	55.5	57.7	58.4	58.3	54.1	54.3
7	Pongola (EU564828)	**66.4**	**67.2**	**67.2**	**69.4**	**65.1**	*83.0*‡		55.5	55.8	57.0	54.1	56.2	55.8	56.2	57.5	55.1	54.5
8	Bunyamwera (NC_001927)	44.1	42.8	45.4	45.0	44.5	45.4	44.5		*90.0*	90.1	87.7	75.8	86.1	*86.4*	*86.1*	57.1	50.6
9	Shokwe (EU564831)	44.1	43.2	45.9	45.0	44.5	45.9	45.0	*97.4*		*87.2*	*85.8*	*74.8*	*87.0*	*87.8*	*84.3*	56.7	50.6
10	Nyando (AM709781)	44.1	43.2	45.9	44.5	44.5	45.4	44.5	92.1	*92.1*		85.9	75.1	85.2	*84.8*	*85.1*	58.4	51.4
11	Ilesha (AM709780)	39.3	38.4	40.6	40.6	39.7	42.8	41.5	87.3	*86.5*	86.5		71.9	81.7	*83.5*	*82.5*	54.2	49.4
12	Germiston (BLCSA)	45.9	45.0	46.7	43.2	43.2	45.9	43.7	74.7	*75.5*	75.5	69.0		75.2	*76.4*	*74.1*	55.4	52.0
13	Cache Valley (X73465)	44.5	43.7	45.9	45.0	44.5	46.7	46.3	90.4	*91.7*	91.7	83.0	74.7		*87.8*	*82.2*	57.4	51.2
14	Fort Sherman (EU564829)	44.5	42.8	46.3	44.1	43.2	47.2	45.4	*93.0*	*92.6*	*92.6*	*84.7*	*75.1*	*96.1*		*82.5*	56.1	50.3
15	Xingu (EU564830)	45.9	45.0	47.2	46.3	46.3	45.0	45.0	*91.3*	*90.4*	*90.4*	*83.0*	*74.2*	*90.0*	*90.8*		57.0	50.7
16	Apeu (DQ188952)	46.3	47.6	46.7	46.7	47.2	41.9	42.8	45.9	47.6	47.6	42.8	47.2	48.0	47.6	47.6		63.0
17	Oropouche (AY237111)	42.4	43.7	42.8	43.2	42.8	43.7	45.0	41.5	41.9	41.9	39.3	41.0	41.5	40.6	39.3	51.5	

*Values in upper right hand and lower left hand corners of quadrant represent nucleotide and amino acid sequence identities, respectively.

†Emboldened values represent highlighted comparisons within the Bwamba and California serogroups.

‡Underlined and italicized values represent highlighted comparisons between viruses of like serogroup.

§ GenBank accession numbers appear next to virus names. Apeu and Oropouche viruses are included as Group C and Simbu serogroup references, respectively.

Phylogenetic analyses. Phylogenetic analyses of newly determined N and NSs gene ORF sequences in comparison to previously determined S segment sequences of selected members of the genus *Orthobunyavirus* were conducted using MEGA version 4 software (Tamura, Dudley, Nei, & Kumar, 2007). Sequences of genomes of the Bunyamwera,

Bwamba, California, Group C and Simbu virus serogroups of the genus *Orthobunyavirus* were aligned using Clustal W (MEGA version 4). Both Neighbor-Joining (NJ) and Maximum Parsimony (MP) trees were generated and analyzed with 2000 replicates for bootstrap testing of each grouping (Satou & Nei, 1987; Eck & Dayhoff, 1966; Felsenstein, 1985). Trees generated by both NJ and MP methods displayed nearly identical topologies with highly comparable bootstrap values for groupings generated from both the N and NSs ORF nucleotide sequences and only the NJ tree generated from the N ORF nucleotide sequences of represented viruses is shown (Figure 5.1). Within the NJ inferred evolutionary history, distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004; Figure 5.1) under the Tamura-Nei model of evolution (Tamura & Nei, 1993).

Results and Discussion

General characteristics of newly derived and previously determined S segment sequences of human pathogens, representing all medically important serogroups of the genus *Orthobunyavirus* for which S segment sequence data have been generated, are presented. S segment sequences of medically important viruses of the California, Bwamba, Bunyamwera, Group C and Simbu serogroups are described (Table 5.1). The total and non-coding region (NCR) nucleotide sequence lengths vary among the compared viruses (Table 5.1). Of the compared viruses, the Afrotropical Bwamba virus has the longest S segment total and 3'NCR sequences, of 1096 and 310 nucleotides, respectively (Table 5.1).

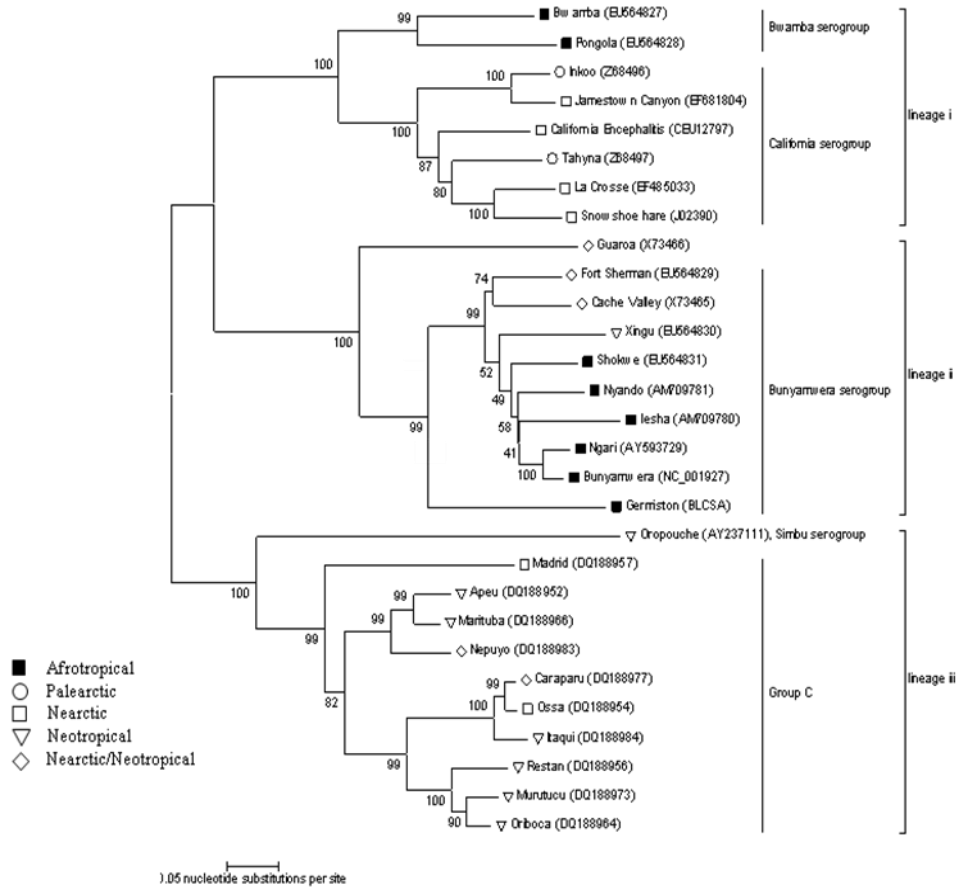

Figure 5.1. Phylogeny of N gene nucleotide ORF sequences of human pathogens of the genus *Orthobunyavirus*. Major biogeographic regions associated with viruses of interest are shown. GenBank accession numbers appear next to virus names.

In addition, the Bwamba serogroup virus S segments have comparatively high A+U contents, with the Bwamba virus S segment having the highest overall A+U content of nearly 65%, which is contributed to by the long A+U rich 3'NCR (Table 5.1). The significance of these findings is not known. The 5' and 3' termini of the newly derived Bwamba and Bunyamwera serogroup S segment sequences are highly conserved, with only the Shokwe and Xingu virus S segments differing from the other presented sequences by the last two nucleotides of the 3' terminal region (AC for CT) and the 3' terminal nucleotide of the S segment (C for T), respectively (data not shown). With regard to coding regions, within the Bunyamwera serogroup, the Nearctic and Neotropical Fort Sherman and Xingu viruses and the Afrotropical Shokwe virus have

identical N and NSs gene ORF nucleotide and amino acid sequence lengths, 702nt/233aa and 306nt/101aa, respectively, to those of the prototype Afrotropical Bunyamwera virus strain. Overall, these N and NSs sequence lengths are generally conserved among represented viruses of the Bunyamwera serogroup (Table 5.1). Of interest, the newly determined N and NSs nucleotide and amino acid sequences of the Afrotropical Bwamba viruses, Bwamba and Pongola, have nearly identical lengths to those of the majority of represented Nearctic California serogroup viruses (Table 5.1). When the coding regions of Bwamba and Pongola viruses are directly compared to the Nearctic California serogroup La Crosse virus strain Human, 1978, only the N gene of Bwamba virus differs by 3 nucleotides (705/708) and one corresponding amino acid (234/235; Table 5.1).

The percent identities among newly derived and previously determined N gene nucleotide ORF and amino acid sequences are presented (Table 5.2). Fort Sherman, Xingu and Shokwe viruses share between 74.1% to 90.0% nucleotide, and from 74.2% to 97.4% amino acid, sequence identities with viruses of the Bunyamwera serogroup (Table 5.2). Among all compared viruses, the Afrotropical Shokwe and Bunyamwera viruses share the highest percent identities of 90.0% and 97.4% for nucleotide and amino acid sequences, respectively (Table 5.2). The Bwamba serogroup viruses, Bwamba and Pongola, share nearly 80% nucleotide and 83% amino acid sequence identities within the serogroup (Table 5.2). Interestingly, the Bwamba serogroup viruses also share, on average, over 70% nucleotide and nearly 67% amino acid sequence identities with the California serogroup reference La Crosse virus, higher percent identities than are shared between any other compared viruses of heterologous serotype (Table 5.2).

Phylogenetic analysis of S segment sequences of selected viruses of the Bunyamwera, Bwamba, California, Group C and Simbu serogroups of the genus *Orthobunyavirus*

depicts three major lineages of orthobunyaviruses with extremely high bootstrap support (99-100%; Figure 5.1). The three major lineages are mostly composed of i) the predominantly Nearctic California serogroup, ii) the predominantly Neotropical Group C viruses and iii) the predominantly Afrotropical Bunyamwera serogroup. As expected, Fort Sherman, Shokwe and Xingu viruses group within the Bunyamwera serogroup of lineage iii with strong bootstrap support (Figure 5.1). However, while the Nearctic, Neotropical and Afrotropical viruses generally cluster with viruses of like geography, there is relatively weak support for some of these groupings, suggesting a limited association between geographic and genetic distinctions among some viruses within the Bunyamwera serogroup (Figure 5.1). Surprisingly, the Afrotropical Bwamba serogroup viruses, Bwamba and Pongola, cluster together in one of two groups represented in lineage i, with the Nearctic and Palearctic California serogroup viruses comprising the second, major lineage i grouping with extreme bootstrap support (Figure 5.1).

Within the Afrotropical biogeographic region, viruses of both the Bwamba and Bunyamwera serogroups of the genus *Orthobunyavirus* have been associated with predominantly febrile human illnesses (Barrett & Shope, 2005; Lanciotti & Tsai, 2007; LinksKalunda et al., 1985; Lutwama et al., 2002; Smithburn et al., 1941). While there is some evidence of Tahyna virus circulation in Africa (Chastel et al., 1983; Kuniholm et al., 2006), all other human pathogens of the California serogroup, and the primarily encephalitic human illnesses associated with these viruses, are known to exclusively occur in Nearctic and Palearctic biogeographic regions (Barrett & Shope, 2005; Gonzales-Scarano & Nathanson, 1996; Lanciotti & Tsai, 2007). The presented molecular and phylogenetic characterizations of the N and NSs ORFs suggest that a common S segment ancestor is shared between the serologically, clinically and geographically

disparate Bwamba and California groups of the genus *Orthobunyavirus*. We speculate that ancestral subpopulations, of unidentified origin, became isolated through host migration and geographic separation of unknown mechanism(s). As a function of this isolation, varying selective pressures, among ecologically distinct vertebrate and arthropod hosts, are proposed to have driven the divergent evolution of these viruses, resulting in the unique and defining clinical and serological characteristics of the California and Bwamba serogroups. We hypothesize that as a consequence of this divergence, the M segment sequences of Bwamba serogroup viruses will likely be more variable than the presented S segment sequences when compared to those of the California serogroup viruses because of the association between M segment encoded glycoproteins and vertebrate and arthropod host cell attachment and entry (Barrett & Shope, 2005; Ludwig et al., 1991), although no information regarding these segments is available.

CHAPTER 6

MULTI-SEGMENT MOLECULAR AND PHYLOGENETIC ANALYSES REVEAL NO EVIDENCE OF S and M GENOMIC SEGMENT REASSORTMENT AMONG DISTINCT VIRUSES OF THE CALIFORNIA SEROGROUP OF THE FAMILY *BUNYAVIRIDAE*

Abstract

To support comprehensive analyses of arthropod-borne viruses of the family *Bunyaviridae*, nucleotide sequences have been determined for multiple segments of viruses of interest (Chapter 2; Lambert & Lanciotti, 2008; Lambert & Lanciotti, 2009). Utilizing these data, in combination with those that were previously generated by others, we have conducted analyses of multiple genomic segments of all viruses currently classified within the California serogroup of the genus *Orthobunyavirus* of the family *Bunyaviridae.*The results of these analyses reveal no evidence of S and M genomic segment reassortment among compared viruses, strongly suggesting that genetic drift is the central driving evolutionary force of the medically important California serogroup.

Introduction

The ability of a bunyavirus co-infection to result in the generation of a novel progeny virus that is the recipient of a combination of genomic segments from serologically related heterologous parental viruses has been well documented for viruses of the genus *Orthobunyavirus* in the laboratory setting (Beaty et al., 1985; Beaty et al., 1997; Borucki et al., 1999; Cheng et al., 1999). In recent years, the description of this phenomenon has extended beyond the laboratory into natural populations, with surprisingly high rates of reassortment reported for orthobunyaviruses of closely related serological classification (Nunes et al., 2005; Reese et al., 2008). These reports include the description of frequent genomic reassortment between viruses of the Group C serogroup of the Brazilian Amazon (Nunes et al., 2005) as well as an approximated 25% genomic reassortment rate reported for variant strains of La Crosse virus of the California serogroup that cocirculate in Wisconsin and Minnesota within the United States (Reese et al., 2008). Additionally, within the Bunyamwera serogroup of the genus *Orthobunyavirus* there is a recent report of segment reassortment within Main Drain and Potosi viruses that can both be vectored by *Aedes* species mosquitoes and occur within small rodent hosts in North America (Briese et al., 2007).

Outstanding among the descriptions of segment reassortment in nature is the discovery of a reassortant orthobunyavirus, Ngari virus that was associated with outbreaks of hemorrhagic human illness that occurred in Africa in the late 1990s (Bowen et al., 2001; Gerrard et al., 2004; Briese et al., 2006). Ngari virus is a member of the Bunyamwera serogroup that possesses S and L segments from an ancestor of Bunyamwera virus; a virus that has been historically associated with outbreaks of mild febrile illness in Africa, and an M segment from Batai virus; a widely distributed virus that has never been associated with human illness. The implications of this finding are

unsettling, with particular regard to the fact that the reassortant Ngari virus is associated with a much more severe human clinical outcome than that of either of its two serologically related parental viruses. This observation indicates that a bunyavirus reassortment event can lead to a dramatically altered spectrum of human disease as has been seen with influenza A viruses (Webster et al., 1992). Despite this discovery, the contribution of segment reassortment to bunyavirus evolution and emergence on broad taxonomic, temporal and geographic scales has been historically defied by a limited availability of nucleotide sequence data for viruses of the family *Bunyaviridae*.

Notable among groups of bunyaviruses that have historically lacked comprehensive molecular description, the California serogroup of the genus *Orthobunyavirus* is of unique interest. Occuring predominantly throughout the Northern Hemisphere, human pathogens of the California serogroup include the previously described La Crosse virus (Chapter 4) and are generally characterized by an encephalitic clinical outcome that can be age associated. Described pathogens also include Snowshoe hare virus that is distributed throughout regions of Eastern Europe and North America and has been associated with encephalitis in children in Canada (Fauvel et al, 1980) and Tahyna virus that is associated with a predominantly febrile human illness (Bardos et al., 1975) throughout Eastern Europe. Also classified within the California serogroup, Inkoo virus, originally isolated from mosquitoes in Finland in 1964 (Brummer-Korvenkontio et al., 1973), has been anecdotally associated with CNS disease in young people in that country and is now known to circulate widely throughout Europe (Putkuri et al., 2007). Additionally, within North America the prototype species of the California serogroup, California Encephalitis virus occurs in the western United States and like other described human pathogens of the serogroup has been associated with encephalitic human illness

(Hammon et al., 1952). Also occuring throughout the United States, Jamestown Canyon virus has similarly been implicated as a cause of encephalitis in predominantly adult populations (Grimstad et al., 1986). In addition to human pathogens, the California serogroup also includes Chatanga, Lumbo, South River, Jerry Slough, San Angelo, Melao, Keystone, Trivittatus and San Angelo viruses that are collectively globally distributed and have not been associated with human illness.

While members of the California serogroup possess distinct associations with specific vector and vertebrate hosts (Beaty & Calisher, 1989), the known geographic distribution of the California serogroup includes regions of shared geography for many of these viruses, as described above. A shared geography allows for the possibility of a segment reassortment event in nature. However, while two other key serogroups of the genus *Orthobunyavirus*, the Group C and Bunyamwera serogroups, have been subjected to multi-segment analyses (Bowen et al., 2001; Gerrard et al., 2004; Briese et al., 2006; Briese et al., 2007; Nunes et al., 2005) to infer segment reassortment in the evolution of their members, no analogous study of the California serogroup has been as-of-yet conducted due to a lack of comprehensive nucleotide sequence data.

To address deficiencies in the understanding of arthropod-borne viruses of the family *Bunyaviridae*, we have generated nucleotide sequence data for multiple genomic segments of many viruses of the *Orthobunyavirus* and *Phlebovirus* genera of the family *Bunyaviridae* (Chapter 2; Lambert & Lanciotti, 2008; Lambert & Lanciotti, 2009). Furthermore, in an attempt to identify the role of segment reassortment in the evolution of the California serogroup of the genus *Orthobunyavirus* we have conducted molecular and phylogenetic analyses of newly derived and previously determined S, M and L genomic segment nucleotide sequences for all known members of this group. Prior to this

study and in light of data presented by others (Bowen et al., 2001; Gerrard et al., 2004; Briese et al., 2006; Briese et al., 2007; Nunes et al., 2005; Reese et al., 2008), we have defined the occurrence of strongly supported, inconsistent groupings of S, M and L segment sequences of an individual virus within divergent clades of phylogenetic trees as evidence of segment reassortment. Furthermore, and also in light of data presented by others (Bowen et al., 2001; Gerrard et al., 2004; Briese et al., 2006; Briese et al., 2007; Nunes et al., 2005) we consider markedly divergent patterns of shared nucleotide sequence percentage identities among compared viruses in the S, M and L segments as additional evidence of segment reassortment.

The Study

Confirmation and alignment of sequences. Previously determined California serogroup N ORF (S segment), M ORF (M segment) and newly determined partial polymerase ORF (L segment; Chapter 2) nucleotide and amino acid sequences were aligned using the Megalign software (DNA*) Clustal W function (Thompson et al, 1994). Prior to alignment, all newly derived L segment sequences were confirmed by methods described in Chapter 2. Additionally, all S, M and L genomic segment sequences that were previously submitted to GenBank by others were additionally confirmed through partial amplification and sequencing using methods described in Chapter 2. Unfortunately, all described efforts for the generation of partial L segment sequences (Chapter 2) from Keystone, Melao, San Angelo and Lumbo viruses failed to generate nucleotide sequence data for comparison. As such, these viruses were not subjected to L segment analyses.

P**hylogenetic analyses and nucleotide and amino acid percentage identity calculations.** Phylogenetic analyses of newly determined sequences in the context of previously determined S, M and L segment sequences of all described viruses that are classified within the California serogroup were conducted using MEGA version 4 software (Tamura, Dudley, Nei, & Kumar, 2007). Both neighbor-joining (NJ) and maximum parsimony (MP) trees were generated and analyzed with 1000 replicates for bootstrap testing (Satou & Nei, 1987; Eck & Dayhoff, 1966; Felsenstein, 1985). Neighbor-joining trees were generated utilizing the Maximum Composite Liklihood method under the Tamura-Nei model of evolution (Tamura & Nei, 1993; Tamura et al., 2004). Percentage identities were calculated through the use of the MEGA version 4 software pairwise distance calculation function using a "p-distance" model (Tamura, Dudley, Nei, & Kumar, 2007).

Results

Phylogenetic trees generated from the full length ORFs of the S and M segments of all individual viruses described within the California serogroup are presented (Figures 6.1 &6.2). Of relevance to the discussion of genomic segment reassortment, all phylogenies generated for the S and M segment ORFs possess nearly identical topologies with similarly strong bootstrap support values for all major groupings (Figures 6.1 & 6.2), indicating a lack of S and M genomic segment reassortment among compared viruses. Unfortunately, both NJ and MP analyses of the L segment resulted in bootstrap support values of less than "50% for nearly all major groupings, indicating very little phylogenetically informative signal within this segment. Therefore, no tree depicting an L segment phylogeny is presented.

Nucleotide and amino acid percentage identities calculated for S, M and L genomic segment sequences of all viruses classified within the California serogroup are presented (Tables 6.1-6.3). For the S segment, nucleotide percentage identities range from 72.1% shared between Trivittatus and Lumbo viruses to 92.6% shared between Jerry Slough and Inkoo and Jamestwon Canyon and Inkoo viruses (Table 6.1). The overall mean percentage identity for all compared California group S segment nucleotide sequences is 80.4% (Table 6.1). For the M segment, nucleotide sequence percentage identities vary from 65.4% shared between Trivittatus and San Angelo viruses to 84.7% shared between Jamestown Canyon and Inkoo viruses (Table 6.2). Additionally, indicating less conservation relative to the S segment, the overall mean percentage identity for all compared California group M segment nucleotide sequences is 70.9% (Table 6.2). For the L segment, nucleotide sequence percentage identities range from 70.5% shared between Tahyna and Jerry Slough viruses to 85.3% shared between Jamestown Canyon and Inkoo viruses (Table 6.3). The overall mean percentage identity for all compared California group L segment nucleotide sequences is 76.5% (Table 6.3), indicating an intermediate level of conservation relative to the S and M segment sequences of compared viruses. Of note, California serogroup amino acid percentage identities are generally higher and reflect the same pattern of sequence diversity as their respective nucleotide sequence percentage identities (Tables 6.1-6.3).

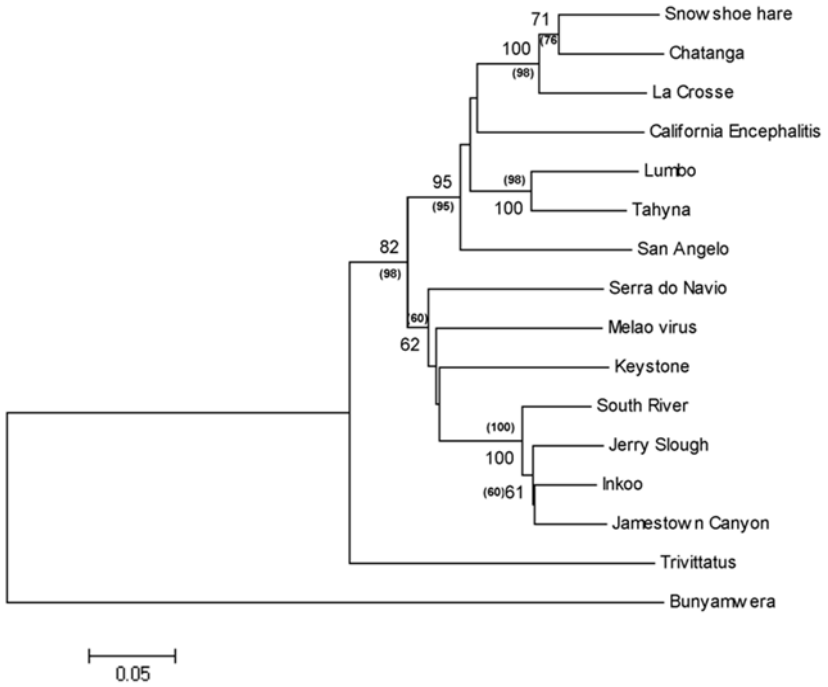

Figure 6.1. California serogoup S segment phylogeny. Both maximum parsimony and neighbor-joining methods were utilized to derive phylogenies from full length N ORF sequences (702 nts). A neighbor-joining tree is presented and corresponding bootstrap values of greater than 60 generated by maximum parsimony and neighbor-joining methods appear within and outside of parentheses, respectively. The scale bar represents the number of nucleotide substitutions per site.

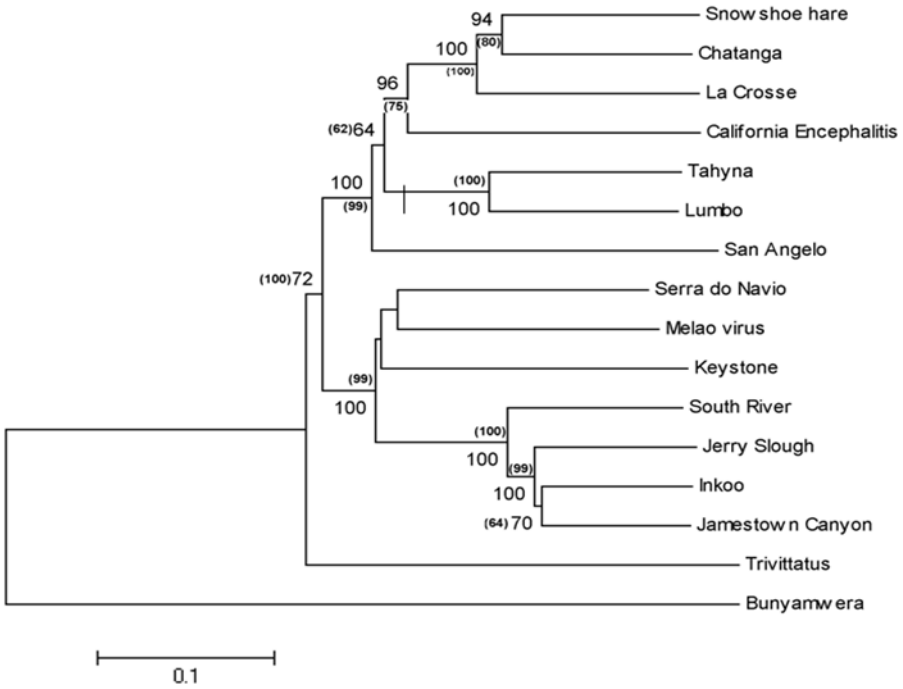

Figure 6.2. California serogoup M segment phylogeny. Both maximum parsimony and neighbor-joining methods were utilized to derive phylogenies from full length polyprotein ORF sequences (4338 nts) of all described viruses of the California serogroup. A neighbor joining tree is shown. Corresponding bootstrap values of greater than 60 generated by MP and NJ analyses appear within and outside of parentheses, respectively.

Table 6.1. S segment N ORF nucleotide and amino acid sequence identities compared among all described human pathogens and related members of the California serogroup of the genus *Orthobunyavirus* shown above and below the diagonal, respectively.

Virus		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
SSH (J02390	1		88.9	88.9	82.1	82.1	82.1	83	77.2	77.8	78.3	78.6	77.4	77.8	78.2	73.5	53
Chatanga (GQ330486)	2	93.9		86.6	81.9	82.6	82.1	82.3	77.2	78.2	78.2	78.1	78.5	78.2	77.8	73.4	53
La Crosse (EU485033)	3	90.5	90		83.2	83.5	82.3	81.5	77.6	78.2	77.4	80.2	78.5	79.8	79.5	73.6	54
CE (CEU127	4	88.7	88.7	85.7		83.6	81.8	82.8	77.4	78.1	79.1	79.3	77.9	79.3	79.1	73.2	55.3
Tahyna (Z68497)	5	88.3	88.7	85.7	90		88.9	82.1	78.2	80.6	79.6	80.3	79.8	79.9	79.3	72.2	55
Lumbo (X73468)	6	87.9	88.7	84.8	90.5	95.7		81.9	78.3	80.1	80.1	80.2	79.8	79.2	78.6	72.1	53.6
SA (U4714	7	88.7	89.6	86.6	89.2	88.7	90		78.8	80.3	79.3	79.5	79.8	80.1	79.2	72.5	54.3
SDN (U4714(	8	80.5	82.7	80.5	84.4	81.8	84	82.7		82.2	81.8	82	82.6	83.6	82.1	74.6	55.6
Melao (MVU12802)	9	82.3	84.4	83.5	83.1	84	83.1	84.4	87.9		82.8	83.5	83.2	82.9	83.8	73.8	54.6
Keystone (U12801)	10	81	83.1	80.5	84.8	82.7	84	84.8	88.7	87.4		83.6	83.3	82.3	82.9	73.6	54.8
SR (U4714(	11	82.7	86.7	83.1	86.1	835	85.7	85.3	91.8	88.7	88.3		91.2	92.9	91.9	75.1	54.4
JS (U127	12	81.4	84.8	82.7	84.8	82.7	84.8	84.4	80.9	87.9	87.9	98.3		92.6	92	74.5	54.3
Inkoo (Z68496	13	82.3	85.3	82.7	85.7	83.1	85.3	84	81.3	87.9	87.9	98.3	97.8		92.6	73.4	54
JC (EF6818(	14	81.4	84	81.4	84	81.8	84	84.4	90	87	87.9	97.4	97.8	97		73.5	53.8
Trivittatus (U12803)	15	73.6	73.6	72.7	74.5	72.7	74.2	73.2	76.2	75.3	74.9	75.7	75.3	75.3	75.3		53.8
Bunyamwera (NC_001927)	16	45.9	45.5	44.2	43.3	45	44.2	45.5	44.2	46.8	46.3	45.5	45.5	45.5	45.5	44.2	

A) Aligned sequences are are 702/231 nucleotides/amino acids in length.
B) CE, California encephalitis; JC, Jamestown Canyon; JS, Jerry Slough; SDN, Serra do Navio; SSH, Snowshoe hare; SR, South River; SA, San Angelo.
C) Highlighted values demonstrate consistently shared patterns of relatively high shared nucleotide sequence percentage identities that are common to the S, M and L segment identity matrices.
D) Bunyamwera virus sequence included as a Bunyamwera serogroup reference.

Table **6.2**. Full length M segment ORF nucleotide and amino acid sequence identities compared among all described human pathogens and related members of the California serogroup of the genus *Orthobunyavirus* shown above and below the diagonal, respectively.

Virus		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
SSH (EU2625	1		89	79.4	73.5	71.2	73	70.6	69.1	69.9	68.5	69.7	69.1	69.7	69.5	65.9	53
Chatanga (EU621834	2	92.2		78	73.8	72.5	73.1	71.8	71.1	69.8	68.7	69.5	69.4	69.5	69.4	66.9	53
La Crosse (NC_00410	3	90.3	88.3		73.7	72.5	72.8	70.8	69.4	69.9	69	68	68.8	68.6	69.5	66.3	52.7
CE (AF123	4	82.1	82.8	81.5		73	73	71.3	69.3	70.1	67.8	68.8	68.5	68.6	68.3	66	53
Tahyna (GQ386841)	5	81.3	82.3	80.4	80.5		81.1	71.2	70.1	70.7	70.3	69.6	69.5	69.6	69.1	67.3	53.6
Lumbo (AF123484	6	78.3	79.7	77.7	78.9	90.1		70.7	69.2	69.8	69.3	69.1	68.9	69.3	68.9	67.5	53.5
SA (AF1234	7	76	76.5	74	76.8	76.3	74		69.1	69	68.1	68.6	67.9	67.7	67.9	65.4	52.7
SDN (AF1234	8	72.9	73.4	72.2	72	72.3	72.2	71		76.3	74.3	74.5	72.9	73.5	73.7	67.8	54
Melao (MVU08805	9	74.7	74.2	74.3	73	74.2	74.3	71.4	84.3		73.3	73.2	73.3	73.5	73.3	67.8	53.9
Keystone (AF123489)	10	72.4	72.7	73.3	71.8	72.3	71.6	70.5	74	79.3		72.7	71.8	71.9	72.2	66.7	53.3
SR (AF1234	11	73.2	73	71.9	71.4	73	71.9	71.1	72.2	79.3	76.1		81.9	82	82.3	66.5	53.7
JS (AF123	12	72.7	72.8	72.8	70.9	72.5	71.4	71.4	78.3	78.4	75.1	93.3		84	83.6	65.9	55
Inkoo (U8805	13	73.8	74.2	72.7	72	73.7	72.5	71.9	71.9	79	75.8	92.7	92.8		84.7	65.7	52.9
JC (HM007	14	73.5	73.6	73.7	71.4	73.2	72.1	71.6	90.1	79.1	76.1	94.5	94.5	94.2		66.1	55
Trivittatus (AF123491)	15	67.6	67.7	67.4	67.1	68	66.9	65.6	66.9	66.2	66.6	66	65.7	65.6	65.7		52.4
Bunyamwera (NC_001926)	16	42.7	42	42.8	43.3	42.2	41.9	42.9	41.9	43.3	42.7	42.9	42.5	42.2	42.5	41.4	

A) Aligned sequences are are 4338/1428 nucleotides/amino acids in length.
B) CE, California encephalitis; JC, Jamestown Canyon; JS, Jerry Slough; SDN, Serra do Navio; SSH, Snowshoe hare; SR, South River; SA, San Angelo.
C) Highlighted values demonstrate consistently shared patterns of relatively high shared nucleotide sequence percentage identities that are common to the S, M and L segment identity matrices.
D) Bunyamwera virus sequence included as a Bunyamwera serogroup reference.

Table 6.3. Partial L segment ORF nucleotide and amino acid sequence identities compared among all described human pathogens and related members of the California serogroup of the genus *Orthobunyavirus* shown above and below the diagonal, respectively.

		1	2	3	4	5	6	7	8	9	10	11	12
SSH (EU20367	1		81.6	81.2	79.3	79.8	75.1	73.3	76.2	75.1	72.8	73.3	54.2
Chatanga (EU616903	2	95		80.5	77.9	79.3	74.9	73	74.4	73.7	72.1	74.2	55.6
La Crosse (NC_00410	3	91.5	90.1		75.8	77.9	74.9	75.3	73	75.3	72.8	75.3	56.5
CE*	4	85.1	83.7	86.5		75.3	77.7	72.8	76.7	74.2	79.8	72.3	58.4
Tahyna*	5	89.4	90.1	87.9	83.7		74.9	72.1	70.5	71.4	71.9	73.6	54.7
SDN*	6	85.1	84.4	82.3	84.4	80.9		75.8	78.6	77.7	75.8	74.7	56
SR*	7	82.3	81.6	79.4	78.7	78.7	83.7		81.2	82.3	84.3	74	55.8
JS*	8	82.3	80.9	80.1	79.4	78.7	87.9	85.1		81.9	80.9	76	58.8
Inkoo (EU78957	9	83	83	80.1	78.7	76.6	85.8	83.7	95		85.3	74.2	56
JC*	10	83	83.7	79.4	76.6	79.4	85.8	80.9	93.6	95		73.3	56.7
Trivittatus*	11	85.9	84.4	83	85.1	80.9	88.7	82.3	85.1	83	81.6		52.6
Bunyamwera (NC_001923)	12	51.8	51.8	52.4	50.4	48.9	50.4	48.9	48.9	49.6	49.6	47.5	

A) Aligned sequences are are 430/141 nucleotides/amino acids in length.
B) CE, California encephalitis; JC, Jamestown Canyon; JS, Jerry Slough; SDN, Serra do Navio; SSH, Snowshoe hare; SR, South River; SA, San Angelo.
C) Highlighted values demonstrate consistently shared patterns of relatively high shared nucleotide sequence percentage identities that are common to the S, M and L segment identity matrices.
D) Genbank accession numbers shown in parentheses only for virus sequences that were previously determined.
E) *Newly determined sequence.
F) Bunyamwera virus sequence included as a Bunyamwera serogroup reference.

serogroup reference, Bunyamwera virus amino acid percentage identities are consistently lower than their respective nucleotide sequence counterparts, indicating a divergent pattern of evolution occuring between Bunyamwera virus and the California serogroup viruses (Tables 6.1-6.3). Finally, of relevance to the pursuit of evidence of segment reassortment, the presented percentage identities calculated for S, M and L genomic segment nucleotide sequences show nearly identical patterns with no virus parings demonstrating marked shifts in relative nucleotide sequence percentage identities (Table 6.1-6.3). These findings provide additional evidence of a lack of genomic segment reassortment among compared viruses.

Discussion

Taken together, the results of our analyses (Figures 6.1 & 6.2; Tables 6.1-6.3) provide no evidence of S and M genomic segment reassortment within and among compared

viruses of the California serogroup of the genus *Orthobunyavirus.* Unfortunately, definitive results from a phylogenetically informative and comprehensive analysis of the L segment preclude the determination of an absence of L genomic segment reassortment among compared viruses. However, the presented L segment percentage identities are generally consistent with those of the S and M segments (Tables 6.1-6.3), suggesting a lack of L segment reassortment among compared viruses.

From an epidemiological perspective our findings are particularly compelling. The absence of evidence of S and M genomic segment reassortment among all compared viruses of the California serogroup makes the potential future discovery and description of such a reassortant virus novel and possibly predictive of an altered clinical outcome. This fact is of unique interest in the southeastern United States, where there is evidence that the paradigm of California serogroup virus ecology and distribution is changing (Gerhardt et al., 2001; Lambert et al, 2010). Within this region of the United States, Keystone, Jamestown Canyon, La Crosse and San Angelo viruses have been known to occur. It is proposed that a genomic reassortment event among heterologous viruses might be facilated by the expansion of *Aedes albopictus* throughout the United States as this vector has well described catholic feeding habits, transovarial transmission capabilities (Tesh and Shroyer, 1980), and the previously mentioned laboratory-capability for the generation of reassortant California serogroup viruses (Cheng et al., 1999).

CHAPTER 7

CONCLUSIONS AND FUTURE DIRECTIONS OF RESEARCH

Conclusions

The genomic characterization, detection and molecular evolutionary analyses of medically relevant and serologically related viruses of the arthropod-borne genera of the family *Bunyaviridae* are described (Chapters 2, 3, 4, 5 & 6). Supporting current and future research and detection efforts, the described efforts have culminated in the successful determination of nearly 50 full length and partial primary nucleotide sequences for greater than 30 viruses of interest (Chapter 2; Lambert & Lanciotti, 2008; Lambert & Lanciotti, 2009). Utilizing these data in support of design and characterization, a molecular consensus assay that utilizes a novel multiplex amplification and sequencing method for the detection and identification of 47 viruses of divergent classification within the *Orthobunyavirus*, *Phlebovirus* and *Nairovirus* genera has been developed (Chapter 3; Lambert & Lanciotti, 2008). Demonstrating the extraordinary utility of newly developed methods (Chapters 2 & 3; Lambert & Lanciotti, 2008; Lambert & Lanciotti, 2009) the discoveries of Tahyna virus in China in 2006, La Crosse virus in *Aedes albopictus* mosquitoes in Dallas County, Texas in 2009 and Toscana viral RNA within the CSF of an American traveler who returned from a trip to Sicily in 2009 are reported (Chapter 4; Kay et al., 2010; Lambert et al., 2010; Lu et al., 2009). Of note and to the best of our knowledge, the detection of La Crosse viral S, M, and L genomic segment RNAs in *Aedes albopictus* mosquitoes that were collected in Texas in the summer of 2009 (Chapter 4; Lambert et al., 2010) marks the first time a molecular assay has been used to simultaneously detect all three segments of the bunyavirus genome in the diagnostic setting. Furthermore and of evolutionary interest, utilizing primary sequence data determined as described in Chapter 2, we have uncovered

a unique and unexpected genetic similarity shared between the serologically, clinically and geographically disparate Bwamba and California groups of the genus *Orthobunyavirus* (Chapter 5; Lambert & Lanciotti, 2009). Finally, utilizing comprehensive nucleotide sequence data that were generated by us and others we have determined that S and M genomic segment reassortment among distinct viruses is not a central driving evolutionary force of the California serogroup (Chapter 6). Despite these accomplishments, we have encountered considerable challenges to the comprehensive molecular understanding of viruses of the family *Bunyaviridae.* The most significant of these continues to be the extraordinary number of highly diverse species classified within the family *Bunyaviridae.* We consider the inability to amplify certain viruses of interest (Chapters 2, 3 & 6) a function of this great diversity, manifested in target nucleic acid heterogeneity. In an attempt to address this challenge and to facilitate a more comprehensive understanding of the *Bunyaviridae*, the following future directions of research are proposed.

Future directions of research

The pursuit of alternative methods for the generation of nucleotide sequence data. Most fundamental to the ability to characterize unknown viral nucleic acids are the strengths and limitations of the tools we use to evaluate these agents. Current molecular methodologies, including those described here (Chapters 2&3) are almost always exclusively dependent on target-specific oligonucleotide priming events for amplification of viral nucleic acid. While these methods confer a high degree of specificity and sensitivity, they are limited in their ability to only prime target-specific reactions. Particularly compelling for the characterization of bunyaviruses, is the advent of sequence independent amplification and sequencing techniques, commonly known as

"shotgun sequencing", for the determination of primary sequence data in the absence of *a priori* knowledge of conserved nucleotide sequences. Prior to the beginning of this study, we attempted sequence independent amplification and sequencing of bunyaviruses, with limited success (data not shown). From these attempts, it was found that an excess of contaminating nucleic acids from host cells, precluded the amplification and sequencing of target bunyaviral nucleic acids (data not shown). With this understanding, we would like to make additional attempts at refining a "shotgun sequencing" method for the determination of additional S, M and L segment primary sequence data from arthropod-borne buyaviruses of interest. Future attempts will incorporate virus purification and enzymatic treatments designed to select for viral particles and nucleic acids, prior to amplification, to enhance the likelihood of generating viral primary nucleotide sequence from randomly amplified fragments. Beyond the scope of our research, the generation of these data is foundational to the future molecular and evolutionary understanding of bunyaviruses as a group and will support the development of additional assays for their detection.

Development of a multi-segment assay. Integrating newly derived sequences into our design, the expansion of the presented molecular consensus assay (Chapter 3; Lambert & Lanciotti, 2009) is proposed. The design of the expanded assay will integrate a capacity for the detection of reassortant viruses by including additional genomic segments as targets for amplification. From what is known about the relatively frequent M-segment driven pattern of bunyavirus genomic segment reassortment in nature (Briese et al., 2006; Briese et al., 2007; Nunes et al., 2005), we propose that the expanded strategy would minimally include S and M genomic segment targets for amplification and detection. This would allow for the detection of the majority of potential reassortant

viruses that could arise from targeted agents. Of interest, the utility of such an assay has been preliminarily demonstrated through the use of a multi-segment pilot strategy for the detection of La Crosse viral S, M and L segment RNAs (Chapter 4; Lambert et al., 2010). At this point in time, the expanded consensus format is envisioned to utilize a similar RT-PCR based strategy for the amplification of multiple fragments from target viruses of interest (Chapter 3; Lambert & Lanciotti, 2009). As presented here (Chapter 3; Lambert & Lanciotti, 2009), species level identification will be facilitated through multiplex sequencing and NCBI BLAST analysis.

Comprehensive evolutionary analyses at the serogroup and genus levels. As a complement to the discussed expansion of our molecular characterization and detection efforts, additional and more comprehensive evolutionary analyses at both the serogroup and genus levels are proposed. Of immediate interest and hopefully including a more comprehensive set of L segment data generated by sequence independent methods, additional analyses of the California serogroup are planned. These analyses will include the investigation of codon sites undergoing positive selection, more refined analyses of nucleotide sequence diversity within and among the S, M and L genomic segments and the determination of putative sites of glycosylation (M segment). In addition, we have generated data from multiple genomic segments of viruses from diverse serogroups of both the *Orthobunyavirus* and *Phlebovirus* genera that have yet to be explored at an evolutionary level (Chapter 2; Appendix 1). Of immediate interest are multi-segment evolutionary analyses within and among serogroups at the *Orthobunyavirus* genus level. Such analyses will facilitate a better understanding of the relative contributions of segment reassortment and genetic drift to the evolution of viruses within different serogroups of the genus *Orthobunyavirus*.

Molecular epidemiology of La Crosse virus. Inspired by the discovery of La Crosse virus in *Aedes albopictus* mosquitoes in Dallas County, Texas in the summer of 2009 (Lambert et al., 2010), we propose additional investigations to facilitate a better understanding of the molecular epidemiology of this important pediatric encephalitic agent. The proposed study will involve the full length genomic sequencing of La Crosse virus isolates of varied geographic, temporal and host sources. Following genomic sequencing, multi-segment analyses will be performed in an attempt to identify the selection of a limited range of M segment genotypes in the human host as has been suggested through limited investigation (Huang et al., 1997) and empirical observation (Lambert et al., 2005). In light of data that indicate an approximated 25% rate of segment reassortment among variant strains of La Crosse virus in nature (Reese et al., 2008), our investigations will focus on revealing any potential inconsistencies in the phylogenetic alignments of S, M and L genomic segment sequences of compared La Crosse virus strains from varied hosts. Such analyses would be performed in an attempt to reveal the potential role of La Crosse virus segment reassortment in the emergence of human disease.

Engage enhanced discussion on the molecular taxonomy of bunyaviruses. The molecular taxonomic classification of bunyaviruses has been historically limited due to a paucity of data and the resultant lack of clear standards. It is my hope that the data that we have generated will foster an enhanced discussion on the molecular taxonomy of arthropod-borne viruses of the family *Bunyaviridae*, with a unique emphasis on defining standards for the distinction between virus species and strains. Of special interest, all of

the described viruses of the California serogroup of the genus *Orthobunyavirus* (Chapter 6) are currently classified as individual strains of California encephalitis virus. Clearly, our analyses indicate that each of the viruses within this serogroup is unique at the nucleotide sequence level (Chapter 6) which argues for their classification as individual species of the California serogroup.

REFERENCES

Altschul S.F., Gish W., Miller W., Myers E.W. & Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* **215,** 403-410.

Arikawa J., Schmaljohn A., Dalrymple JM. & Schmaljohn, C.S. (1989). Characterization of Hantaan Virus Envelope Glycoprotein Antigenic Determinants Defined by Monoclonal Antibodies. *J Gen Virol* **70,** 615-624.

Armstrong P.M. & Andreadis T.G. (2006). A new genetic variant of La Crosse virus *(Bunyaviridae)* isolated from New England. *Am J Trop Med Hyg* **75,** 491-6.

Bardos V. & Danielova V. (1959). The Tahyna virus—a virus isolated from mosquitoes in Czechoslovakia. *J Hyg Epidemiol Microbiol Immunol* **3,** 264–76.

Barrett, A.D. & Shope R.E. (2005). *Bunyaviridae,* In *Topley and Wilson's Microbiology and Microbial Infections,* 10th edn, pp. 1025-1058. Edited by B. Mahy & V. Ter Meulen. Washington D.C.: ASM Press.

Beaty, B. J., and C. H. Calisher. 1991. *Bunyaviridae*--Natural history. *Curr Top Microbiol*
Immunol **169,** 27–78.

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

Beaty B. J., Borucki M., Farfan J. & White D. (1997). Arbovirus-vector interactions: determinants of arbovirus evolution, In *Factors in the emergence of arbovirus diseases*, pp. 23-35. Edited by J. F. Saluzzo & B. Dodet. Paris, France: Elsevier.

Bird B.H., Githinji J.W., Macharia J.M., Kasiiti J.L., Muriithi R.M., Gacheru S.G., Musaa J.O., Towner J.S,. Reeder S.A., Oliver J.B., Stevens T.L., Erickson, B.R., Morgan L.T., Khristova M.L., Hartman A.L., Comer J.A., Rollin P.E., Ksiazek T.G. & Nichol S.T. (2008). Multiple virus lineages sharing recent common ancestry were associated with a large Rift Valley fever outbreak among livestock in Kenya during 2006-2007. *J Virol.* **82,** 11152-66**.**

Bishop D. H. (**1996).** Biology and Molecular Biology of Bunyaviruses, In *The Bunyaviridae*, pp. 19-53. Edited by R.M. Elliott. London, England: Plenum Press.

Bishop D. H. & Beaty B.J. (**1988).** Molecular and biochemical studies of the evolution, infection and transmission of insect bunyaviruses. *Philos Trans R Soc Lond Biol Sci* **321,** 463-483.

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

Blakqori G., Delhaye S., Habjan M., Blair C.D., Sanchez-Vargas I. & K.E. Olson. (2007). La Crosse bunyavirus nonstructural protein NSs serves to suppress the type I interferon system of mammalian hosts. *J Virol* **81,** 4991–4999.

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. *J Gen Virol* **80,** 3173-3179.

Bowen, M.D., Trappier S.G., Sanchez A.J., Meyer R.F., Goldsmith C.S., Zaki S.R., Dunster L.M., Peters C.J. , Ksiazek T.J. & Nichol S.T. (2001). A reassortant bunyavirus isolated from acute hemorrhagic fever cases in Kenya and Somalia. *Virology* **291,** 185-190.

Bridgen, A. & Elliott, R. M. (1996). Rescue of a segmented negative-strand RNA virus entirely from cloned complementary DNAs. *PNAS* **93**, 15400-15404.

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. *J Virol* **80**, 5627-30.

Briese T., Kapoor, V. & Lipkin, W.I. (2007). Natural M-segment reassortment in Potosi and Main Drain viruses: implications for the evolution of orthobunyaviruses. *Arch Virol* **152,** 2237-47.

Brummer-Korvenkontio, M. (1973). Arboviruses in Finland. V. Serological survey of antibodies against Inkoo virus (California group) in human, cow, reindeer, and wildlife sera.
Am J Trop Med Hyg **22:**654-661.

Bulychev V.P., Alekseev A.N., Kostiukov M.A., Tukhtaev T.M. & Gordeeva Z.E.(1985). Isolation of Tahyna virus from mosquitoes collected in Dushanbe. *Med Parazitol* **4,** 81–3.

Burney, M.I., Ghafoor A., Saleen M., Webb, P.A. & Casals, J. (1980). **Nosocomial Outbreak of Viral Hemorrhagic Fever Caused by Crimean Hemorrhagic Fever-Congo Virus in Pakistan, January 1976.** *Am J Trop Med Hyg* **29,** 941 - 947.

Calisher C.H., Weinberg A., Muth D.J. & Lazuick J.S. (1987). Toscana virus in United States citizen returning from Italy. *Lancet* **1,** 165–6.

Calisher, C.H., & Karabatsos, N. (1988). Arbovirus serogroups: definition and geographic distribution. In *The Arboviruses: Epidemiology and Ecology*, pp.19-57. Edited by T.P. Monath. Boca Raton, Fl: CRC Press.

Campbell G.L., Mataczynski J.D., Reisdorf E.S., Powell J.W., Martin D.A., Lambert A.J., Haupt T.E., Davis J.P. & Lanciotti R.S. (2006). Second human case of Cache Valley virus disease. *Emerg Infect Dis* **12,** 854-6.

Casals, J. & Whitman L. (1961). Group C, a new serological group of hitherto undescribed arthropod-borne viruses. Immunological studies. *Am J Trop Med Hyg* **10,** 250-258.

Centers for Disease Control. (1986). *Aedes albopictus* introduction-Texas. MMWR **35,** 141-2.

Charrel R.N., Gallian P., Navarro-Mari J.M., et al. (2005). Emergence of Toscana virus in Europe. *Emerg Infect Dis* **11**, 1657–63.

Chastel C., Bach-Hamba D., Launay H., Le Lay G., Hellal H., & Beaucournu J.C. (1983). Arbovirus infections in Tunisia: new serological survey of small wild mammals**.** *Bull Soc Pathol Exot Filiales* **76**, 21-33.

Cheng, L.L., Rodas J., Schultz K.T., Christensen B.M., Yuill T.M. & Israel, B.A. (1999). Potential for evolution of California serogroup bunyaviruses by genome reassortment in *Aedes albopictus*. *Am J Trop Med Hyg* **60,** 430-438.

Crabtree M., Sang R., Lutomiah J., Richardson J. & Miller, B. (2009). Arbovirus Surveillance of Mosquitoes Collected at Sites of Active Rift Valley Fever Virus Transmission: Kenya, 2006–2007. *J Med Entomol* **46,** 961-964.

Daubney, R., Hudson, J.R. & Garnham, P.C. (1931). Enzootic hepatitis or Rift Valley fever. An undescribed virus disease of sheep, cattle and man from East Africa. *J Path Bact* **34,** 799-804.

Di Nicuolo G., Pagliano P., Battisti S., et al. (2005). Toscana virus central nervous system infections in southern Italy. *J Clin Microbiol* **43,** 6186–8.

Dobler G., Treib J., Haass A., Frösner G., Woesner R. & Schimrigk K. (1997). Toscana virus infection in German travellers returning from the Mediterranean. *Infection* **25,** 325.

Eck R.V. & Dayhoff M.O. (**1966**) Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Springs, Maryland.

Eddy, G.A. & Peters, CJ. (1980). The extended horizons of Rift Valley fever: current and projected immunogens. *Prog Clin Biol Res* **47,** 179-191.

Eitrem R., Vene S. & Niklasson B. (1990). Incidence of sand fly fever among Swedish United Nations soldiers on Cyprus during 1985. *Am J Trop Med Hyg* **43,** 207–11.

Elliott, R. M., Bouloy, M., Calisher, C. H., Goldbach, R., Moyer, J. T., Nichol, S. T., Pettersson, R., Plyusnin, A., & Schmaljohn., C. S. (2000). Family *Bunyaviridae,* In *Virus taxonomy: seventh report of the International Committee on Taxonomy of Viruses.* pp. 599-621. Edited by M. H. V. van Regenmortel, C. M. Fauguet, and D. H. L. Bishop. San Diego, Calif.: Academic Press.

Fauvel M., Artsob H., Calisher C.H., Davignon L., Chagnon A., Skvorc-Ranko R. & Belloncik S. (1980). California group encephalitis in three children in Quebec: Clinical and serologic findings. *J Canad Med Assoc* **122,** 60-63.

Felsenstein J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783-791.

Flick, R. & Pettersson, R.F. (2001). Reverse genetics system for Uukuniemi virus (Bunyaviridae): RNA polymerase I-catalysed expression of chimeric viral RNAs, *J Virol* **75,** 1643–1655.

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. Emerg Infect Dis* **7,** 807-11.

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. *J Virol* **78,** 8922-8926.

Grimstad P.R., Calisher C.H., Harroff R.N. & Wentworth B.B. (1986).
Jamestown Canyon virus (California serogroup) is the etiologic agent of widespread infection in Michigan humans. *Am J Trop Med Hyg* 35, 376-86.

Grimstad P.R., Kobayashi J.F., Zhang M.B. & Craig G.B. Jr. (1989). Recently introduced Aedes albopictus in the United States: potential vector of La Crosse virus (*Bunyaviridae:* California serogroup). *J Am Mosq Control Assoc* **5,** 422-7.

Gonzalez-Scarano, F., Pobjecky, N., & Nathanson, N. (1984).
La Crosse bunyavirus can mediate pH-dependent fusion from without. *Virology* 132: 222-225.

Gonzales-Scarano, F. & Nathanson, N. (1996). *Bunyaviridae,* In *Fields' Virology*. pp. 1473-1494. Edited by B. N. Fields, D. M. Knipe, and P. M. Howley. Philadelphia, Pa.: Lippincott Williams & Wilkins.

Gould E.A., Higgs S., Buckley A. & Gritsun T.S. (2006). Potential arbovirus emergence and implications for the United Kingdom. *Emerg Infect Dis* **12,** 549–55.

Habjan M., Pichlmair A., Elliott R.M., Overby A.K., Glatter T., Gstaiger M., Superti-Furga G., Unger H. & F. Weber. (2009). NSs protein of rift valley fever virus induces specific degradation of the double-stranded RNA-dependent protein kinase. *J Virol* **83,** 4365–4375.

Haddow A.D. & Odoi A. (2009). The incidence risk, clustering, and clinical presentation of La Crosse virus infections in the eastern United States, 2003–2007. PLoS ONE 2009; 4: e6145. doi:10.1371/journal.pone.0006145

Hammon, W.M. & Reeves, W.C. (1952). California encephalitis virus, a newly described agent. *Calif Med* **77,** 303-9.

Hammon, W.M., Reeves, W.C. & Sather, G. (1952). California encephalitis virus, a newly described agent. II. Isolations and attempts to identify and characterize the agent. *J Immunol* **69,** 493-510.

Hart, T. J., A. Kohl, and R. M. Elliott. (2009). Role of the NSs protein in the zoonotic capacity of Orthobunyaviruses. *Zoonoses Public Health* **56,** 285-296.

Henderson, W.W., Monroe, M.C., St Jeor, S.C., Thayer, W.P., Rowe, J.E., Peters, C.J. & Nichol, S. T. (1995). Naturally occurring Sin Nombre virus genetic reassortants. *Virology* **214,** 602-610.

Hertig, M. & Sabin, A. (1964). Sandfly fever (Papatasi, Phlebotomus, three day fever). *In Preventative Medicine in World War II.* pp. 109-174. Edited by J. Coates, E. Hoff and P. Hoff. Washington, DC: Office of the Surgeon General.

Honig J., Osborne J. & Nichol, S. (2004). The high genetic variation of viruses of the genus Nairovirus reflects the diversity of their predominant tick hosts. *Virology* **318,** 10-16.

Huang C., Thompson W.H., Karabatsos 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 Res* **48,** 143-8.

Hubálek Z. Zeman P., Halouzka J., Juřicová Z., Šťovíčková E., Bálková H., et al. (2005). Mosquitoborne viruses, Czech Republic, 2002. *Emerg Infect Dis* **11,** 116–8.

Ikegami, T., C. J. Peters, and S. Makino. (2005). Rift Valley fever virus nonstructural protein NSs promotes viral RNA replication and transcription in a minigenome system. *J Virol* **79,** 5606-5615.

Jin H & Elliott RM. (1993). Characterization of Bunyamwera virus S RNA that is transcribed and replicated by the L protein expressed from recombinant vaccinia virus. *J Virol* **67,** 1396–404.

Kay M.K., Gibney K.B., Riedo F.X., Kosoy O.L., Lanciotti R.S. & Lambert, A.J. (2010) Meningoencephalitis Caused By Toscana Virus Infection in an American Traveler Returning from Sicily — 2009. *Emer Inf Dis in press*

Keegan K. & Collett M.S. (1986). Use of bacterial expression cloning to define the amino acid sequences of antigenic determinants on the G2 glycoprotein of Rift Valley fever virus. *J Virol* **58,** 263-270.

Klimas, R.A., Thompson W.H., Calisher, C.H., Clark, G.C., 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 recombination of La Crosse virus. *Am J of Epi* **114,** 112-131.

Kuniholm M.H., Wolfe N.D., Huang C.Y., Mpoudi-Ngole E., Tamoufe U., LeBreton M., Burke D.S. & Gubler D.J. (2006). Seroprevalence and distribution of Flaviviridae, Togaviridae, and Bunyaviridae arboviral infections in rural Cameroonian adults. *Am J Trop Med Hyg* **74**,1078-83.

Kuno, G., Mitchell C.J., Chang G.-J. J. & G. C. Smith. (1996). Detecting bunyaviruses of the Bunyamwera and California serogroups by a PCR technique. *J Clin Microbiol* **34,** 1184-1188.

Kunz C., Buckley S.M., & Casals J. (1964). Antibodies in man against Tahyna and Lumbo viruses determined by hemagglutination-inhibition and tissue culture neutralization tests. *Am J Trop Med* **13**,738-741.

Lambert, A.J., Nasci, R.S., Cropp, B.C., Martin, D. A., Rose, B. C., Russell, B. J. & Lanciotti, R.S. (2005). Nucleic acid amplification assays for detection of La Crosse virus RNA. *J Clin Microbiol* **43**, 1885-1889.

Lambert A.J., and R.S. Lanciotti. (2008). Molecular characterization of medically important viruses of the genus Orthobunyavirus. *J GenVirol* **89,** 2580-2585.

Lambert A.J. & Lanciotti R.S. (2009). Consensus Amplification and Novel Multiplex Sequencing Method for S Segment Species Identification of 47 Viruses of the Orthobunyavirus, Phlebovirus, and Nairovirus Genera of the Family Bunyaviridae. *J Clin Microbiol* **47,** 2398 - 2404.

Lambert A.J., Blair C.D., D'Anton M., Ewing W., Harborth M., Seiferth R., Xiang, J. & Lanciotti, R.S. (2010). La Crosse virus in Aedes albopictus mosquitoes, Texas, USA, 2009. *Emer Inf Dis* **16,** 856-858.

Lanciotti R.S., Kosoy O.L., Laven J.J., Panella A.J., Velez J.O., Lambert A.J. & Campbell G.L. (2007). Chikungunya virus in US travelers returning from India, 2006. *Emerg Infect Dis* **13,**764-7.

Lanciotti R.S., Kosoy O.L., Laven J.J., Velez J.O., Lambert A.J., Johnson A.J., Stanfield S.M. & Duffy, M.R. (2008). Genetic and serologic properties of Zika virus associated with an epidemic, Yap State, Micronesia, 2007. *Emerg Infect Dis* **14,** 1232-1239.

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., Panigrahy 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,** 2333-7.

Lanciotti R.S. & Tsai, T.F. (2007). Arboviruses. In the *Manual of Clinical Microbiology.* pp. 1486-1501. Edited by P.R. Murray, E.J. Baron, J.H. Jorgensen, M.L. Landry & M.F. Pfaller. Washington, D.C.:ASM Press.

Li, D., Schmaljohn A.L., Anderson K. & Schmaljohn C.S. (**1995).** 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.

LinksKalunda M., Lwanga-Ssozi C., Lule M., & Mukuye A. (1985). Isolation of Chikungunya and Pongola viruses from patients in Uganda.
Trans R Soc Trop Med Hyg. **79**,567.

Lu, Z., Xin-Jun, L., Shi-Hong, F., Song, Z., Zhao-Xia, L., Xin-Hua, Y., Yu-Ping, F., Lambert, A.J., Da-Xin, N., Feng-Tian, W., Su-Xiang, T., Nasci, R.S., Yun Feng, Q.D., You-Gang, Z. Xiao-Yan, G., Huan-Yu, W., Qing, T., and Liang, G.D. (2009). Tahyna Virus and Human Infection, China. *Emer Inf Dis* **15,** 306-309.

Lutwama J.J., Rwaguma E.B., Nawanga P.L., & Mukuye A. (2002).
Isolations of Bwamba virus from south central Uganda and north eastern Tanzania. *Afr Health Sci* **2**, 24-8.

Ludwig, G.V., Israel B.A., Christensen B.M., Yuill T.M. & Shultz K.T. (1991). Role of La Crosse virus glycoproteins in attachment of virus to host cells. *Virology* **181,** 564-571.

Lvov DK. Kostiukov MA, Pak TP, Gordeeva ZE & Bun'etbekov AA. (1977).
Isolation of Tahyna virus (California antigenic group, family *Bunyaviridae*) from the blood of febrile patients in the Tadzhik SSR. *Vopr Virusol* **6,** 682–5.

Lyons, M.J. & Heyduk, J. (1973). Aspects of the developmental morphology of California encephalitis virus in cultured vertebrate and arthropod cells and in mouse brain. *Virology* **54**, 37-52.

Mir, M. A. & Panganiban, A. T. (2006). The bunyavirus nucleocapsid protein is an RNA chaperone: possible roles in viral RNA panhandle formation and genome replication. *RNA* **12**, 272-282.

Montgomery, R. (1917). On a tick-borne gastro-enteritis of sheep and goats occurring in British East Africa. *J Comp Pathol* **30,** 28-57.

Moore C.G. & Mitchell C.J. (1997). *Aedes albopictus* in the United States: ten-year presence and public health implications. *Emerg Infect Dis* **3,** 329-34.

Nichol, S. T. (2001). Bunyaviruses. In *Fields' Virology,* p.p.1603-1633. Edited by B. N. Fields, D. M. Knipe, P. M. Howley, & D. E. Griffin. Lippincott Williams & Wilkins: Philadelphia, Pa.

Nichol, S., Beaty, B.J., Elliott, R.M., Goldbach, A., Plyusnin, A., Schmaljohn, C.S. & Tesh, R.B. (2005). Bunyaviridae In Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U. & Ball L.A. (eds) Virus Taxonomy: VIIIth Report of the International Committee on Taxonomy of Viruses, Elsevier Academic Press: Amsterdam.

Novoa, R.R., Calderita, G., Cabezas, P., Elliott, R.M. & Risco, C. (2005). Key Golgi Factors for Structural and Functional Maturation of Bunyamwera Virus *J Virol* **79,** 10852 - 10863.

Nunes M.R., A.P. Travassos da Rosa, S.C. Weaver, R.B. Tesh & Vasconcelos P.F. (2005). Molecular epidemiology of group C viruses (*Bunyaviridae, Orthobunyavirus*) isolated in the Americas. *J Virol* **79,** 10561-70.

Patterson, J. L., B. Holloway, and D. Kolakofsky. (1984). La Crosse virions contain a primer-stimulated RNA polymerase and a methylated cap-dependent endonuclease. *J Virol* **52:**215-222.

Plassmeyer M.L., Soldan S.S., Stachelek K.M., Martín-García J. & González-Scarano F. (2005). California serogroup Gc (G1) glycoprotein is the principal determinant of pH-dependent cell fusion and entry. *Virology* 2005 338: 121-32.

Putkuri, N. Vaheri, A. & Vapalahti, O. (2007). Prevalence and Protein Specificity of Human Antibodies to Inkoo Virus Infection. *Clin and Vac Immun,* **14,** 1555-1562.

Ravkov, E. V., Nichol, S.T. & Compans, R.W. (1997). Polarized entry and release in epithelial cells of Black Creek Canal virus, a New World hantavirus. *J Virol* **71,** 1147–1154.

Reese S., Blitvich B., Blair C., Geske D., Beaty B., & W.C. Black, IV. (2008). Potential for La Crosse virus segment reassortment in nature. *Virol J* **5,** 164: doi:10.1186?1743-422X-5-164.

Saitou N. & Nei M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* **4**,406-425.

Sánchez-Seco M.P., Echevarría J.M., Hernández L., Estévez D., Navarro-Marí J.M., & Tenorio, A. 2003. Detection and identification of Toscana and other phleboviruses by RT-nested-PCR assays with degenerated primers. *J Med Virol* **71,**140-149.

Schmaljohn, C. S. & Nichol, S.T. (**2007).** *Bunyaviridae*: In *Fields' virology*, p.p. 1741-1778. Edited by B. N. Fields, D. M. Knipe, P. M. Howley, & D. E. Griffin. Lippincott Williams & Wilkins: Philadelphia, Pa.

Shope, R. E. & Causey, O.R. (1962). Further studies on the serological relationships of group C arthropod-borne viruses and the application of these relationships to rapid identification of types. *Am J Trop Med* **26**, 189-208.

Smith, J. F. & Pifat, D. Y. (1982). Morphogenesis of sand fly fever virus (Bunyaviridae family). *Virology* **121,** 61-81.

Smithburn, K. C., Mahaffy, A. F., & Paul, J.H. (1941) Bwamba fever and its causative virus. *Am. J. Trop. Med.*, 21, 75-90.

Smithburn, K.C., Haddow, A.J. & Mahaffy, A.F. (1946). A neurotropic virus isolated from *Aedes* mosquitoes caught in the Semliki forest. *Am J Trop Med Hyg* **26,** 189-208.

Sonderegger B., Hachler H., Dobler G. & Frei M. (2009). Imported aseptic meningitis due to Toscana virus acquired on the island of Elba, Italy, August 2008. *Eurosurveillance* **14,** 1–2.

Swanepoel R. (1995). *Nairoviruses infections*, In *Kass handbook of infectious diseases, exotic viral infections.* pp. 285-293. Edited by J. Porterfield. Oxford: Chapman and Hall Medical.

Swanepoel R., Shepherd A.J., Leman P.A., Shepherd S.P. & Miller G.B. (1985). A common-source outbreak of Crimean-Congo haemorrhagic fever on a dairy farm. *S Afr Med J* **68,** 635-7.

Tamura K. & Nei M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* **10,** 512-26.

Tamura, K., Nei, M., & Kumar S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *PNAS* **101**,11030-11035.

Tamura K., Dudley J., Nei M. & Kumar S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* 10.1093/molbev/msm092.

Tesh R.B. & Shroyer D.A. (1980). The mechanism of arbovirus transovarial transmission in mosquitoes: San Angelo virus in *Aedes albopictus. Am J Trop Med Hyg* **29,** 1394-1404.

Thomas, D., Blakqori, G., Wagner, V., Banholzer, M., Kessler, N., Elliott, R.M., Haller, O. & Weber, F. (2004). Inhibition of RNA polymerase II phosphorylation by a viral interferon antagonist. *Journal Of Biol Chem* **279,** 31471-31477.

Thompson J.D., Higgins D,G. & Gibson T.J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22,** 4673-4680.

Venturi G., Ciccozzi M., Montieri S., et al. (2007). Genetic variability of the M genome segment of clinical and environmental Toscana virus strains. *J Gen Virol* **88,** 1288–94.

Venturi G., Madeddu G., Rezza G., et al. (2007). Detection of Toscana virus central nervous system infections in Sardinia Island, Italy. *J Clin Virol* **40,** 90–91.

Verani P., Ciufolini M.G., Nicoletti L., et al. (1982). Ecological and epidemiological studies of Toscana virus, an arbovirus isolated from Phlebotomus. *Ann Inst Super Sanita* **18,** 397–9.

Verani P. & Nicolleti, L. (1995). In *Kass handbook of infectious diseases, exotic viral infections.* pp. 295-317. Edited by J. Porterfield. Chapman & Hall Medical Medical: Oxford, UK.

Vorou R., Pierroutsakos I.N. & Maltezou H.C. (2007). Crimean-Congo hemorrhagic fever. *Curr Opin Infect Dis* **20,** 495-500.

Weaver S.C. & Reisen W.K. (2010). Present and future arboviral threats. *Antiviral Res* **85,** 328–45.

Weber, F., Dunn, E.F., Bridgen, A. & Elliott, R.M. (2001). The Bunyamwera Virus Nonstructural Protein NSs Inhibits Viral RNA Synthesis in a Minireplicon System. *Virology* **281,** 67-74.

Weber, F. Bridgen, A., Fazakerley, J.K., Streitenfeld, H. Randall, R.E. & R.M. Elliott. (2002). Bunyamwera bunyavirus nonstructural protein NSs counteracts the induction of alpha/beta interferon. *J Virol* **76,** 7949–7955.

Webster R., Bean W., Gorman O., Chambers T., & Y. Kawaoka. (1992). Evolution and ecology of influenza A viruses. *Microbiol Mol Biol Rev* **56,** 152-179.

Weekly epidemiological record. (2008). **83,** 157.

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. *Arch Virol* **151,** 2253-60.

Yandoko, E. N., S. Gribaldo, C. Finance, A. Le Faou & B. H. Rihn. (2007). Molecular characterization of African orthobunyaviruses. J. Gen. Virol. 88:1761-1766.

Appendix 1, Nucleotide sequences determined in this study. Viral complementary strand sequences are presented in the 5'-3' direction. For full length S segment sequences, the N and NSs ORFs of presented sequences are underlined and highlighted, respectively. Partial sequences are not annotated. For virus sequences that have been submitted to GenBank, accession numbers appear in parentheses next to the strain designation for each virus.

Fort Sherman, 86MSP18 virus strain, 950 nt; complete S segment (EU564829):

5'AGTAGTGTACTCCACACTAATAACTTACAAATCTTGAGAACCTTTGATACATCCAGTA

CTAGCTTGGCCTTTAAATGATTGAGCTAGAATTTAATGATGTCGCTGCTAACACCAGCAGTAC
TTTTGACCCAGAGGTTGCATACATTAACTTTAAGCGTATCTACACCACTGGGCTTAGTTATGAC
CACATTCGAATCTTCTACATTAAAGGACGCGAGATTAAAACTAGTCTCACAAAAAGAAGTGA
ATGGGAGGTTACGCTTAACCTTGGGGGCTGGAAGGTTACTGTATTTAATACAAATTTTCCTGG
CAACCGGAACAGTCCAGTTCCAGACGATGGTCTTACCCTCCACAGACTCAGTGGATTCCTGCC
CGGTACCTACTTGAGAAGATTCTAAAGGTGAGTGAACCAGAAAAGCTGATTATCAAGTCCAA
AATAATCAACCCTTTGGCTGAAAAAAATGGGATTACATGGTCAGATGGAGAAGAGGTGTACC
TCTCATTCTTCCCAGGATCTGAAATGTTCTTAGGAACATTCAAGTTTTATCCACTGGCAATCGG
AATTTACAAAGTCCAGAGAAAAGAAATGGAACCTAAATACCTGGAAAAAACCATGAGGCAG
AGATACATGGGCTTGGAAGCATCAACATGGACAGTCAGCAAAGTGAATGAGGTTCAGTCTGC
TCTGACGGTAGTTTCAGGACTTGGATGGAAGAAAACAAATGTAAGTGCTGCTGCTAGGGAGTT
CCTAGCAAAGTTTGGAATCAATATGTAATCAGACAGGAGGACAAAATCAGATCAAATTAACT
TATTTCCCAATTTGGCTAAAAGAGTCTTCGGACTCAAACAAACAGCATTTTGGGTGGGTGGTT
GGGGACAGAAAAAAATCGGAATACATCAAGATGAGTTATAATATGCTGTTTAAGTTTTAGGT
GGAGCACACTACT3'

Shokwe, Sa Ar 4042 virus strain, 958 nt; complete S segment (EU564831):

5’AGTAGTGTACTCCACGCTACAAACTTGCCATTGCTGAGAATTACACTGCTATTGAATC

CAACAAAAAGTCATTCAAGCTTTTTGATGATTGAGTTGGAATTCCATGATGTCGCTGCTAACA
CCAGCAGTACTTTTGACCCAGAGGTCGCATACGCTAACTTTAAGCGTGTCCACACCACTGGGC
TTAGTTATGACCACATACGAATCTTCTACATTAAAGGACGCGAGATTAAAACTAGTCTCGCAA
AAAGAAGTGAATGGGAGGTTACACTTAACCTTGGGGGCTGGAAGATTACTGTATATAATACG
AATTTTCCTGGAAACCGGAACAACCCAGTTCCTGACGATGGTCTTACCCTCCACAGACTCAGT
GGATTCCTTGCCAGGTACCTACTTGAGAAAATGCTGAAGGTGAGTGAACCAGAAAAACTGAT
TATAAAATCAAAAATAATCAACCCATTGGCAGAAAAAAATGGAATAACCTGGACTGATGGAG
AGGAAGTTTACCTCTCCTTTTTCCCAGGAGCTGAAATGTTCTTAGGGACCTTTAAATTCTACCC
ATTGGCAATTGGAATCTACAAAGTCCAGCGTAAAGAGATGGAGCCTAAATACCTGGAAAAAA
CAATGAGGCAAAGGTACATGGGCTTGGAAGCAGCAACTTGGACTGTTAGCAAGTTAACTGAA
GTCCAGGCGGCTCTAACAGTTGTTTCAGGCTTAGGATGGAAGAAAACTAACGTCAGTGCAGC
AGCTCGAGATTTCCTTGCTAAGTTTGGAATCAGCATGTAAGCAAGGATTCATTTTCAATCTGG
CAAAATTTCAGTTTTTCAATTTGGCTAAAAGGGTTGTATCAACCCACAAATAACAGCATGATG
GGTGGGTGGTTGGGGACAGAAAGACAGCGGCAAAAATCATTATGCGTCATTATTGGTATTAT
AAGTTTTAGGTGGAGCACACTAAC3’

Wyeomyia, Original virus strain, 1110nt; complete S segment (FJ235921):

5’AGTAGTGTACTCCACACTAATACAAAACTAGAATACTAATAAAAGCCAAGAAAATAA

CTAAATCAAAACATGTCTGAAATTGTATTTGATGATATTGGCCAGATTGTTGCTAGTACTTTCA
ACCCAGATAATCAATATATTGGGTTCAAAAGAACTTTTACAATTGGTCTTTCTTATGATCAAAT
AA*GAATCTTCTTTTTGAATGGGAAAAAGGCTAAGGGACACTTTCAAAAAGAAGTGAGGAAA
CTGTCACTTTGAACTTTGGTGGCTGGAGGGTTCCTATCATTAATACTCACTTTCCGGGGAATCG
AAATATGGACCTTGCAGATGATGCTCTCACTCTGCATAGAGTAAGTGGCTACCTGGCTAGATA
CCTGCTTGAGAAGATCTTGAGCGCCCAGGAACCTGAGAAAGTAATTATCAAGACTAAAATAG
TCAACCCAATTGCTGCATCTAATGGGATTTCTTGGGATGATGGCTATGAAGTCTATCTCTCTTT
CTTTCCAGGGACAGAGATGTTCCTTGAAGCTTTTAAATTCTACCCTCTAGCCATAGGGATTTAC
AAAGTTCAAAAAGGGATGATGGATGTGAAATTTTTGGAAAAAACCATGAGACAAAAATATGC
TGGATTGGATGCAACCGTTTGGACTCAGCAGAAATATACAGATGTCATTAATGCCCTGCTTGT
GGTGGTGGCCTTGGATGGAAAAAATCTAATGTGAGTGCTGCAGCCAAAGATTTCCTTTCCAAG
TTTGGGATCCAAATTTAAGTAAAATGGTCAACTCATAAAATAAATCCAAATCATGATATAAAT
AAAAAATACAAAAACAAATAAAAATATATATAAAACAAAAAAACACAAAAAAACACATAAA
AAAACACAAAAACACAAAACAAAAAATAAAAATAAAAACAAAAAACACAAAAAAAACCCCC
AAAAGCAAACCCTAAACCCCCATCAATACAACAGCATGTAAAATGGGTGGGTGGTAGGGGAA
CAAGTAATGCTGCGTAATTAATTTTGCAGAATCTAAAATAATCTTATATGTATGACTATCTTAT
TTTATACATTTTTTACTTTTTAAAGTATTAGGTGGAGCACACTACT3’

*truncated NSs ORF

Xingu, BeH 388464 virus strain, 942nt; complete S segment (EU564830):

5’AGTAGTGTACTCCACACTACATACTTGCCATTCATAAAGTGAGAACTTATCAATCAAT

GCTAGAGTTGGAATTCAATGATGTCGCTGCTAACACCAGCAGTACTTTTGACCCAGAGGTCGG
ATATGCTAATTTTAAGCGTATCCACACCAACGGGCTTAGTTACGACAACATACGAATCTTCTA
CATTAAAGGACGCGAGATTAAAACTAGTCTCGCAAAGAGAAGTGAGTGGGAGGTTACACTTA
ACCTTGGGGGCTGGAAGATTACTGTATACAATACGAATTTTCCTGGCAACCGGAACCACCCAA
TTCCTGACGACGGTCTTACCATCCACCGACTCAGTGGATACTTGGCCAGGTATTTACTTGAGA
AAATATTAAAAGTCAGTGATCCTGAAAAATTGCTTATTAAAACAAAAATTATCAATCCACTGG
CAGAAAAAAATGGGATCACATGGAATGATGGTGAAGAAGTCTATCTTTCATTCTTCCCGGGAT
CTGAGATGTTTCTTGGAACTTTCAAATTTTATCCTCTCGCAATTGGAATTTACAAGGTTCAGCG
GAAAGAAATGGAACCCAAATATCTTGAAAAAGCAATGCGTCAACGCTACATGGGATTAGAGT
CAGCTACATGGACTGTCAGTAAGCTGAATGAAGTGCAATCAGCACTTACTGTTGTATCTGGCT
TAGGTTGGAAGAAAACAAATGTGAGTGCTGCTGCTAGAGAGTTTTTGGCTAAATTTGGCATAA
ACATGTAAATGTTATGCCATGGTGCCTGTAATTAGATCAAAATGTTCATTTGTTAATTTGGCTA
AAAGAGTCTTCGGACTCAACAAAAAACAGCATAAAGAAAGGGTGGGTGGTTGGGGACAGAA
ATATCAAAGCATGCTTTTTCATTAATGCACATCTGTCTCTAATAAGTTGTAGGTGGAGCACACT
ACC3’

Bwamba, M459 virus strain, 1096nt; complete S segment (EU564827):

5’AGTAGTGTACTCCACTTAAGTACTAAAGCTATTAATTGAGGACATCGACTAATA

AAAGCAACTTTAAGTAGAGGCACTATTATGTCCGATTTGATCTTCTATGATGTCGACTCG
GTTAATGCCAATGAGTTTGACCCTGATACAGGATACCTGGATTTTAAAAATAACTATCC
AGGGGCACTCAATACAAATACCGCTAGGACATTCTTCCTCAATGCCGCAAAGGCCAAG
AATGTGCTCCGCAATAAACCTGACAAGAAGGTCAATCCTAAATTTGGAAACTGGGAGG
TGGAGGTTGTCAATAATCATTTTCCTGGAAACAGGAACAACCCAATTGGTAAAGATGAT
CTTACCCTCCACAGAATTTCAGGATATTTAGCAAGATGGATTTTGGAGGAATATAAGAG
AGATGATGAATCAGAAAAGGAAATCATTGAAAGCACAGTTATTAATCCAATAGCTGAA
TCAAATGGGATTAGATGGAGTGATGGAGCAGAGATATATCTATCATTTTTCCCTGGCAC
AGAGATGTTTTTGGAACCATTCAAATTCTATCCACTAGCAATTGGCATATACAGAGTCA
AACATAAGATGATGGATGCACAGTTCCTGAAAAAAGCCTTGCGGCAGAGATATGGGAA
AATGACAGCTGAAAAATGGATGTCAACAAAAGTGAAGGCTATAGGTGAGGCTGTTAGA
AATGTAGAAAAGCTAAAGTGGGGCAAGGGTGGTTTGAGCGATGCTGCAAGAAACTTCC
TAAAGAAGTTTGACATTGCAATGATATAATCAGCCATATGCAAATATTTAGCTATTACT
GAAACATATAACCAATATCAATAACCATAGGACATATATAAAATAAAATATAAAATAC
AAAAAAACAAATAAAACATATATAAAAATAATAAAAAATATAGGGTTAGGGTAAAAAC
AAAAACAAAAAACACAAAAAAACACACAAAAACACAAAAAAACAAAAAAACACAAC
ACACACACAACAGCAATAAAGAAAGTAGTGAACAGTGTTCTCATGGCCGATATTAATCT
TTGCTTGGCTAATGGAAATAATTAAAGGTACTTAGTGGAGCACACTACT
3’

Pongola, Sa Ar 1 virus strain, 921 nt; complete S segment (EU564828):

5’AGTAGTGTACTCCACGAGAATACTTAAAGCCGTTAACAAAATATCTAGATAGCTGAAAGCA
TTCTAAGTGGACTGCACTAATATGTCGGGAGATCTAGTTTTTTATGATGTCGACTC

CGTCAATTCCAATGCGTTTGATCCTGATGCAGGATATGTGGCATTTTCAAACAACCATCC

AGGGGCGCTTAATATCACTACAGCTAGGACATTTTTCCTCAATGCCGCAAAGGCCAAGGGTGT
CCTCCGTAATAAACCTGACAAGAAGGTTAAACCTAAATTTGGAGACTGGGAGGTGGAGATTG
TCAATAATCATTTTCCTGGAAACAGGAACAATACAATTGGTGAAGACGATCTTACCCTCCACA
GGATTTCAGGATATTTAGCTCGTTGGGTTCTTGAAGAACATTCTAAAGATGATGAAAGTGAGA
AAGAATTAATTGAGAGTACAATTGTGAACCCCATTGCTGAGTCAAATGGGATAAAGTGGAGC
GATGGTGCAGCAATTTACTTATCATTCTTCCCAGGAACTGAAATGTTTCTTGAACCGTTTAAGT
TCTTTCCACTTGCAATTGGAATCTACCGCGTTAAGCATAAAATGATGGATGCCCAATTCCTTAA
GAAGGCCTTGCGTCAAAGATATGGGAAACTTACAGCAGAAAAGTGGATGGCAACAAAGGTCA
AATCAATTGGAGAAGCTGTTAGAGCTGTAGAGAAACTGAAGTGGGGAAAAGGAGGAATTAGT
GATGCAGCCAGGACTTTCCTTAAGAAATTTGATATTGCACTAATCTAAACCATTATATTGAAT
CAAAATTCCCAAATTTATATACAACAAATTGGCTATTCCAAAAGGACTTCGGTCCACAAACTC
AGCACAACTTGTTTATTAGACATTAAAAAATGTATTCCGTGGAGCACACTACT3’

Lokern, FMS 4332 virus strain, 655nt; S segment partial N gene:

5’ACTAAAAACTTGCGAACTTTGATAAATCTAAAAATTCCGGAAGACATAACCCAGTCTTTTAA
TGATTGAATTAGAATTTCATGATGTCGCTGCTAACTCCAGCAGTACTTTTGACCCAGAGGTCG
CATACGCTAGTTTTAAGCGTGTCCACACCACTGGGCTTAGTTACGACCACATTCGAATCTTCTA
CATTAAAGGACGAGAGATTAAAACTAGTCTCTCAAAGAGAAGTGAATGGGAGGTTACGCTTA
ACCTTGGGGGCTGGAAGGTTGCTGTATTTAATACAAATTTTCCTGGCAACCGGAACAGTCCAG
TTCCAGACGATGGTCTTACCCTCCACAGACTCAGTGGATTCCTTGCCAGGTACCTACTTGAGA
AAATATTGAAAGTGAGTGAACCAGAGAAGCTCCTGATCAAATCAAAAATCATAAATCCACTT
GCAGAAAAAAATGGCATAACATGGGCAGATGGCGAAGAGGTTTACCTCTCATTCTTTCCAGG
CTCAGAGATGTTTTTGGGCATATTCAAATTCTACCCATTGGCTATTGGTATTTACAAGGTTCAG
CGTAAAGAAATGGAACCAAAATACCTGGAAAAAACAATGCGTCAGAAGTATATGAACATGGA
TGCCGCTACATGG
ACTGTGACCCAAGTC3’

Playas, 75V3066 virus strain, 718nt; S segment partial N gene:

5’AGTAGTGTACTCCACACTAATAACTTTCAAACTTTGGTAAACTTGAGAATTGCAAGGTGTAA
CTTGGCATTCTAATGATTGAGTTGGAATTTAATGATGTCGCTGCTAACACCAGCAGTACTTTTG
ACCCAGAGATTGCATACGTTAACTTTAAGCGTATCCACACCACTGGGCTTAGTTATGACCACA
TTCGAATCTTCTACATTAAAGGACGCGAGATTAAAACTAGTCTCACAAAAAGAAGTGAATGG
GAGGTTACGCTTAACCTTGGGGGCTGGAAGGTTACTGTATTTAATACAAATTTTCCTGGCAAC
AGGAACAGTCCAGTTCCAGACGATGGTCTTACCCTCCACAGACTCAGTGGATACCTTGCCAGG
TACCTACTTGAGAAAATACTGAAGGTGAGTGACCCAGAGAAAGTCATTATAAAATCAAAGAT
AATAAATCCACTAGCTGAAAAGAATGGAATAACCTGGTCTGATGGCGAAGAGGTTTACCTCTC
TTTCTTTCCTGGCTCAGAGATGTTCTTAGGAACATTTAAATTTTACCCTCTGGCAATAGGAATT
TACAAAGTTCAGAGGAAAGAGATGGAACCTAAGTACTTGGAGAAGACAATGCGCCAGAGGTA
TATGGGGCTAGAAGCGTCAACCTGGACAATCAGTAAAGTCAATGAAGTTCAGGCTGCCCTTAC
AGTGGTTTCTGGGCTAGGCTGGAAGAA3’

Santa Rosa M2-1493, virus strain, 301nt; S segment partial N gene:

5’CTTTCAATATTTTCTCAAGTAGGAACCTGGCAAGGAATCCACTGAGCCTGTGGAGGGTAAGA
CCATCGTCTGGAACTGGACTGTTCCGGTTGCCAGGAAAATTTGTATTAAATACAGCAACCTTC
CAGCCCCCAAGGTTAAGCGTAACCTCCCATTCACTTCTCTTTGAGAGACTAGTTTTAATCTCTC
GTCCTTTAATGTAGAAGATTCGAATGTGGTCGTAACTAAGCCCAGTGGTGTGGACACGCTTAA
AACTAGCGTATGCGACCTCTGGGTCAAAAGTACTGCTGGTGTTAGCAGA3’

Alenquer, BeH 301101 virus strain, 354nt; S segment partial N gene (FJ235922):

5’CGGCAAAGCTGGGATGCATCATCGGTCTTGGATAATTTTTAGAGAGACCATCCATGTGTCTT
CCTGTTACAGGCAAGTATTCTTCCACCACCTCAGTAGCTTGGCAAGTGTAACCAGCAAGGGCA
GCTGTCACTCTAGATAGGGTCAGATCATCCCTGGATGGGTTCCCTTCTTTCAGGTTGTATTTCT
TTATAAGCCCTTGAACAATTTTCTTCCCTTTATCAGACATCTTTGTTATCATCTTGCTGGGCTTG
TTGCCCCTAGTGAGACATAAGATGATCATCTGCTTTGCATCATCCTTCCAGTGGTTCCCCCCTC
TGATCTGCAACAGCTCAATAACCCTACGAGCATCAA3’

Candiru, BeH 22511 virus strain, 355nt; S segment partial N gene (FJ235923):

5’CGGCAAAGCTGGGATGCATCATCGCCCTGGGGTAGTTCTTGGATAGATCATCCATGTTCTTC
CCAGTGACAGGCAGAAACTCTTCAACATATTCTGTGGCCTGGCAGGTGTATCCAGCTAAAGCA
GCAGTCACTCTAGACAGTGTCAAATCATCTCTAGAAGGGTTGCCCTCCTTGAGAGAATATCTC
TTGACAAGGGCTTGAACAATTTTCTTGCCCTTCTCTGACATCTTCACCATCATCTTGGAGGGCT
TGTTTCCTCTTGTAAGGCACAGAATAATCATCTGCTTAGCATCTTCAACCCAGTCATCTCCACC
TCTCTCTACCAGCAACTCCATAACCCTCTTAGCATCAAA3’

Chagres, JW 10 virus strain, 338nt; S segment partial N gene (FJ235924):

5’GCTTCCCAGGTACTGAGAATCTTGGCTGAGAGAGGTGGAGCTAACTGGCGTGAGGATGCCA
AGAAAATGATTATTCTCTCTCTCACTAGGGGCAATAAACCCTCTAAGATGATGGAGAGAATGT
CTGAGGCTGGAAAGGCAACTGTGTCGTCGCTAATGACTCGCTACAACCTTAAATCTGGAAACC
CCGGGAGGAATGATCTGACTCTGTCTAGAGTGGCAACTGCCCTTGCTGGATGGACATGTCAGG
CAATTCCAATTGTCCAAGACTTTCTTCCAGTTACTGGTGTAGTGATGGACGGTCTCTCTCCTGC
TTACCCGAGGCCGATGATGCACCC3’

Cacao, VP 437R virus strain, 359nt; S segment partial N gene (FJ235925):

5’GCAAAGCTGGGATGCATCATTTGTCTGGGATAATTGTTTGACAACAAGTCCATTGATTTCCC
AGTTACCGGTAGATAATCTTGTACTATGGGTGCTGCCTGGCATGTCCAGCCTGCTAGTGCAGT
AGCTATCCTAGAGAGGGTTAGATCCTTTCTTCCAGGGTTGCCAGACTTTAGATTGTATCTCTTC
ACCAATGCGTTCACGGTGGACTTCCCTGCATCAGACATTCTTTCTGTCATCTTACTGGGCTTAT
TGCCTCGAGTGAGCGATAGAATTATCATCTTCCTTGCATCCTGTTTCCAGTCATTCTTGCCTTTC
TCTACTAAGGCCTTAATGACTTCACCAGCATCAAATCCTTG3’

Itaporanga, Original virus strain, 369nt; S segment partial N gene (FJ235926):

5’TTTTGCTTATCAAGGATTTGATGCTAGAACCGTTGTTAGGCTGGTCACTGAGAGAGGAGGTG
GAAACTGGCAAGAGGATGTTAAGAGACTCATCGTCTTGGCAGCCACTCGGGGAAATAGGGTC
AACAAGATGAAGGCCAAGATGTCAGACAAGGGTAAAGGAGAACTTACAAAGCTCGTCAGCA
CCTACAAGCTAAAGGAGGGAAATCCTGGCAGAGAGGACTTAACTCTGTCAAGAGTAGCATGT
GCTTTTGCCACCTGGACATGCACTTGCATTACTTCAATTGCTGAGTATCTGCCAGTAACTGGCT
CCACCATGGACAGTTACAGTCCTAATTATCCAAGGACCATGATGCATCCCAGCTTTGC3’

Gabek Forest, SudAn 754-61 virus strain, 145nt; S segment partial N gene (FJ235927):

5’AAGTATAAGCTAAAGTCAGGAAACCCTGGCAGGAATGACTTGACTCTATCTAGAGTTGCTG
CTGCTTTTGCCTCCTGGACTTGCAATGCCATCTATCATGTGCAGTACTACCTCCCGGTGACTGG
AAACCACATGGATGCCATCA3’

Turuna Bear, 352492 virus strain, 362nt; S segment partial N gene (FJ235928):

5’CAAAGCTGGGATGCATCATGGGCCTAGGATAGCTTGGGGACAGGCTATCCATGTGACGGCC
AGTCACAGGCAGATACTCCTCAACCACTTCCGTGGCTTGACAGGTGTATCCTGCGAGTGCTGC
AGTCACTCGAGAAAGGGTTAAATCATCCCTAGATGGGTTCCCTTCTTTCAGGTTGTATTTCTTT
ATCAGCCCTTGCACAATCTTCTTCCCTTTCTCTGACATTTTGGCCATCATTTTTGTTGGCTTGTT
TCCTCTTGTCAAGCACAAGATAATCATTTGCTTGGCGTCTTCCATCCAATTGTCACCCCCTCTG
AGCTTGAGGAGTTCAATGACTCTACGAGCATCAAATCCTTGATAA3’

Rio Grande, TBM3-204 virus strain, 314nt; S segment partial N gene (FJ235929):

5’TGCTCCATTAACAGGCATATAGTTCTCAACGGCCTCTGTGGCTTGACATGTCCAAGGAGCAA
AGGCAGCAGCAATCCTAGCCAGAGTTAGATCATCCCTTCCTGGAGATGATGACTTTAGCTGGT
AAGTTTTGACCAGATCATCCAGAACAGCAATACCCTCTGGGGACATCTTCTTCCTCATTTTGTC
AGGCTTATTGCCCCTCGTCAAAGCAAGAATTATCATCCTTTTCACATCCTCTTGCCACTTTCTC
CCCTTGGCCCTCTCTTGCACGAGTTCAACTACTCGGTTTGCATCAAATCCTTGATAAGCAA3’

Nique, 9C virus strain, 366nt; S segment partial N gene (FJ235920):

5’AGTCCAGCAAAGCTGGGATGCATCATGGCCCTTGGATAATTTTTGGACAATCCATCCATGTG
CCTCCCGGTAACAGGTAAAAACTCTTCTACAACCTCGGTTGCTTGACAAGTGTAGCCGGCTAA
TGCAGCTGTTATCCTTGACAAAGTGAGATCATCCCTAGAAGGGTTTCCTTCTTTAAGGTTGTAT
TTCTTTATGAGCCCCTGGACAATCTTCTTCCCTTTGTCTGACATCTTAGTCATCATCTTTGTTGG
CTTGTTACCACGTGTTAAGCATAAGATTATCATTTGCTTAGCGTCCTCCATCCAGGCATCACCT
CCGCGAGACTTCAAGAGTTCAATCACCCTTCTTGCATCAAATCCTTGA3’

Santa Rosa, M2-1493 virus strain, 450nt; M segment partial polyprotein gene:

5’TCTGATCTAGTGTAACTGATGCTTTTGTCTTGAACCATCCAGAGCTAAGGGTTGTACCTGAC
ACTTCAAAATGATTTAGTTTATCTGTCTGGAAGATTATCTGGGCCTCCTCTTTGTCAATAGTTA
TGGTACAATCTCTGCTACAAACATATGTTTTTGTCATCAGACTAAGGTCAGAGTTCACTTCTAT
GACATTTATACTCCCACCGGCTGTTTGAACTGGGTTACACAATTTCCAATCTTTTATTATCCAG
TTTCGGTATACCTTATTGTGTGCAAAAATTCCTGTTTCATTCTTTTTATAAATTATCTCTGATTT
CAACATCGAGACATCATCCTTAAGGCAATATTCTGATATAGCAGTTTTCGACCTGGTCTCAGC
TATTAATTGTCCTCCATGGAAACACCTTGTTATGATTGGACTCCCTGCTGAGATTCCAATAATC
ATG3’

Playas, 75V3066 virus strain, 511nt; M segment partial polyprotein gene:

5’AGGTTCTATCCAGGGTGACAGACGCCTTGCTTTTAAACCATCCTGAACTTAATGTAGTGCCA
GACACTTCAAAATGATTTAGCTTATCTGTTTGGAAGATAATTTGAGCTTCCTCTTTATCTATGG
TTATAGTGCAATCTCTACTACATATGTATGTCTTTGTCATTAGACTAAGATCTTTGTTAACCTC
AATTACATTTATTGTTCCTCCTGCAGTAGGTATAGGGTTGCAGAGTTTCCAATCTTTTATTGTC
CAGCTTCTAAAAACTTTATTATGGCCAAATAATCCAGTATCATTCTTTATGTATACCACTTCAG
ATTTTAACATTGAGACATCATCCTTGAGGCAGAATTCTGATAAGATGTAGCAGACTTTGACTC
TGCTATTAATTGGCCCCCATGGAAACACCTAGTTGCAATAGGGTTGCTTGTTGAGAACCCTAA
AAAGGCTAGTATGAGCAGGAATAGCATTTTGAAGAACCTTTATCTAGATGTGGATTGTATCGG
TAG3’

Lokern, FMS 4332 virus strain, 446nt; M Segment partial polyprotein gene:

5’TCTCAGCAGGGAGTCCAATCATAACAAGGTGTTTCCATGGAGGACAATTAATAGCTGAGAC
CAGGTCGGAAACTGCTATATCAGAATATTGCCTTAAGGATGATGTCTCGATGTTGAAATCAGA
GATAATTTATAAAAAGAATGAAACAGGAATTTTTGCACACAATAAGGTATACCGAAACTGGA
TAATAAAAGATTGGAAATTGTGCAACCCAGTTCAAACAGCCGGTGGGAGTATAAATGTCATA
GAAGTGAACTCTGACCTCAGTCTGATGACAAAAACATATGTTTGTAGCAGGGATTGTACCATA
ACTATTGACAAAGAGGAGGCCCAGATAATCTTCCAGACAGATAAACTAAATCATTTTGAAGT
GTCAGGTACAACCCTTAGCTCTGGATGGTTCAAGACAAAAGCATCAGTTACACTAGATCAGAC
TGCGAACACA3’

Alenquer, BeH 301101 virus strain, 570nt; M segment partial polyprotein gene:

5’TTCCTAAGAGACTGCGCAGCCCTTTTATGTGGATATTTATGCTACTCTGTTGGTTAATCCAGC
TGGTAAAGAAAGGGTTGAGATCAATGAGCCTCAGAATAAACAATTCCATTGGGTGGACCAAT
CATGCAGAATTGCAAGAGGTGATTAATCACAGGCC
CATCGCTCAGAGACGACCCATCCCTAGATTCCAGGCCACTATGTTCATCCTATTCTCTATTTTT
TCCTTAGGGCTTTCTTGCTCAGAAACCACCCTTTCCAATTCAAAACAGACAAAGTGTGTTCAGT
CAGGAGGGAATGTGAAGTGCACAATAAGTGCA
ACCATTACTATGAAGGCTGGTATTATAGGAGGGGAATCTTGTTTTATAATCAAAGGGCCAATG
GACAACCAACAGAAGACTGTAAGAATAAAAACTGTTTCAAGTGAAATCGTGTGTAGAGAAGG
AAATAGCTTCTGGACAAGCCATTATACCCCATACT
GTCTGAGCTCTAGAAGGTGCCACCTGGTTGGTGATTGCACTGGAAATAGATGTCAAAGCTGGA
CTGATGATTTGGTGTCAACTGAATTCA3’

Candiru, BeH 22511 virus strain, 687nt; M segment partial polyprotein gene:

5’AAGCTCTCTTATAGTGTCAATCCTGTTATCAATTATCATATATTTTGTTTTGATCATAATTGG
CAAGATGTTATACTTCTTCAAAGTCATCCCAAAGAAGTTGAGAAGTCCTTTCATGTGGCTAGC
CCTATTATTTAGATGGGTGATTGACTTGGTCAGGAGCACTTTATCCAAGCTCAAGACCAGACT
TAATAGATCTATTGGCTGGGATGGTCAAGATAGAGGCAATGAACTGAGGGAAGTCAGAGTGA
GGAGGCCCATCCCTAGGTTTGCCGCAACTCTGTCCATAATTTGCTTGCTCACCACTGGTGCCCT
GGCCTGTTCAGAAACGTTAATCTCGAATTCAAAACAAACCAAATGTGTGCAGTCATCTGATAA
GGTGAAATGCTCTGTGTCAGC
CACGGTGACCTTAAAAGCTGGCATTATTGGTGCTGAGTCATGTTTTATTTTGAAGGGTCCAAT
GGAGAATCAGAGAAAGACAATACGTATAAAAACATTATCAAGCGAGCTAGTGTGTAGAGAGG
GCAATAGCTTCTGGACTAGTCACTATAGTCCCACATGTTTGAGTTCTCGCAGATGCCACCTGAT
GAGTGACTGTACTGGCAATAGATGCCAGAGATGGACAGATGAAGAGGTTTCTCAAGAGTTTA
AAGGAGTCAATGACAATATGGTTATGAATGAAAA3’

Nyando, MP401 virus strain, 487nt; L segment partial polymerase gene:

5’CTAGGATAAGATCAGCAAGGGATGCCACAGTGGCAAAAGATATTGACACAGATCTATTGAT
GTCAAGGCATGATTACTTTGGTCAGCAAGTCTGTCGAGCTTTAGGCATTGAGTATAGAAATGA
TATCCCGCTTATTGATATATTCTTGGAAATTGATTCGGACTTTGACCCGTTATCAATGTCAATA
CCAGCTATCACTCCAGACAACTACCAGTTTGTAGATAACATCTTATTCATCATTGATTACAAA
GTATCTGTTTCAAAAGATAGCTCAATCCAAACATTCAATAAATACACAGAATTGATAGGCCCC
ATTTGTAAGAATATAGGCATCATGTTGGATGTGGTGATTATCCGAGCGAATCCCCTAACTCAG
GAGGTGTCAATCAATTCAGATTTGTTCAAGAACTATTTTACAATCAGAAATCTAGAGTTTGAT
TTCCACCAATTTGCTGACTTAAAAAGAATACTATATGAGAAATTTTC3’

Guaroa, CoH 352111 virus strain, 439nt; L segment partial polymerase gene:

5’TCAAACTCTAGATTTCTGATTGTAAAATAGTTCTTGAACAAATCTGAATTGATTGACACCTC
CTGAGTTAGGGGATTCGCTCGGATAATCACCACATCCAACATGATGCCTATATTCTTACAAAT
GGGGCCTATCAATTCTGTGTATTTATTGAATGTTTGGATTGAGCTATCTTTTGAAACAGATACT
TTGTAATCAATGATGAATAAGATGTTATCTACAAACTGGTAGTTGTCTGGAGTGATAGCTGGT
ATTGACATTGATAACGGGTCAAAGTCCGAATCAATTTCCAAGAATATATCAATAAGCGGGATA
TCATTTCTATACTCAATGCCTAAAGCTCGACAGACTTGCTGACCAAAGTAATCATGCCTTGAC
ATCAATAGATCTGTGTCAATATCTTTTGCCACTGTGGCATCCCTTGCTGATCTTATCCTAG3’

Trivittatus, Eklund virus strain, 430nt; L segment partial polymerase gene:

5’AAGCAAGATGGAACCTGGTGAATATCAACAATTTTTAGCACGCATAAACAGTGCAAAAGAT
GCCTGCATTGCTAAAGATATTGATGTTGACCTACTTATGGCCAGACATGATTACTTCGGCAAA
GAATTGTGTAAGGCACTAAATATAGAATATAGAAATGATGTTCCATTTGTAGATATAATTTTA
GATATAAGACCAGATTTTGATCCACTATCTATAGAATTACCACATATCACACCTGACAATTAC
TTATACATCAATAATATTTTGTATATCATAGATTACAAAGTATCTGTATCAAACGAGAGTAGT
GTTATTACGAATACCAAATATTACGAAATGACCAGAGATATATCTGACCATTTAAATATACCC
ATAGAGATAGTCATAATCAGAATA
GACCCAATTAGCAGAGAATTGCACATTAGC3’

Tahyna, Bardos 92 virus strain, 560nt; L segment partial polymerase gene:

5’ATCTACAAAACTTACAAAAACTTTAGTTTGAAACCAATATATGTATCATGGACCATCAGGAA
TACCAGCAGTTCCTCGCGAGGATCAATACTGCAAGGGATGCCTGCATTGCGAAGGATATAGA
CGTGGACCTGTTGATGGCTAGGCATGACTATTTCGGAAAGGAGTTATGCAAATCTTTAAATAT
AGAATATAGGAATGATGTACCATTTATAGATATAATCCTAGATATAAGACCTGAAACCGACCC
TTTAACTGTAGATGCCCCTCATATTACTCCTGATAATTATCTGTATGTAGACAACATATTGTAT
CTAATAGATTATAAAGTATCCGTATCCAATGAGAGTAGCATAATAACATATGATAAATACTAT
GAACTAACTAGAGATATATCTAGACGAATGAATATAGAAATAGAGATAGTTATAGTGAGAAT
AGATCCTGTAAGTAAAGAGTTGCATATTAGCTCCAATAGATTCAAAGAATTGTATCCGGCACT
AGTGGTTGATATAAACTTTGATCAATTCTTTGATTTAAAACAATTGTTGTATGAGAAA3’

Jerry Slough, BFS 4471 virus strain, 483nt; L segment partial polymerase gene:

5’TTCTTCAAACGAATATGGATAATACAGAATATCAACAATTCCTTGCGCGTATCAATTCAGCT
AAAGATGCATGTGTTGCTAAAGACATTGATGTTGATTTACTAATGGCAAGACATGACTACTTT
GGCAGAGAATTATGTAAATCACTAAACATAGAATATAGAAATGATGTCCCATTTGTGGATATC
CTGTTAGATATAAGACCTGATATAGACCCTCTAACACTGGAAATTCCTCATATAACACCAGAC
AATTACCTATATGTAGACAATATTTTATATATTATAGATTATAAGGTTTCTGTTTCTAACGAGA
GTAGTGTTATAACTAATACAAAGTATTTTGAAATGACTAGAGACATTGCATCTATATTAAACC
TACATATAGAAATAGTTATTATTAGAATAGACCCTATTAGTCGCGATCTATATATCAGTTCAG
ATAGGTTTAAGGATTTATTTCCTACATTAGTAGTTGACATTA3’

South River, NJ0-94F virus strain, 560nt; L segment partial polymerase gene:

5’TATTTACAAAACTTACAAATACCAATAATCACTTTAGAGTGACTATGGATAACACAGAATAT
CAACAATTTCTAGCTCGCATTAATGCAGCAAGGGATGCATGTGTCGCAAAAGATATTGATGTG
GACTTATTAATGGCAAGACATGATTATTTTGGGAAAGAACTTTGCAAATCGTTAAATATAGAA
TACAGGAATGACGTTCCTTTTGTAGATATATTATTAGATATAAAGCCAGATATAGACCCCTTG
ACATTAGAAATTCCACATATAACCCCAGATAATTATCTATACTTAGACAACGTTTTATATATA
ATAGACTATAAGGTTTCTGTATCCAATGAGAGTAGTATCATAACTAATACAAAGTATTTTGAG
ATGACTAGAGATATAGCTTCTGCATTAAACCTACACATTGAGGTAGTAATTATCAGGATAGAT
CCAATTAGTCGTGATTTACATATTAGCTCAGACAGATTTAAGGATTTATTTCCCACACTAGTAG
TAGATATCAATTTTAATCAATTTTTTGATTTAAAACAGGCTTTATATGAGAAGTTCGGT3’

Serra do Navio, BeAr 103645 virus strain, 552nt; L segment partial polymerase gene:

5’TCATATAACAATTGCTTTAAATCAAAAAATTGGTTAAAATTTATATCTACAACTAATGTAGG
GAATAAGTCTTTGAATCTATCAGAACTAATATATAATTCTCTACTTATAGGATCTATTCTAATG
ATAACTATTTCTATAGGTATGTTAAGCTTATCTGAAATATCTCTTGTTAATTCAAAATATTTAG
TATTTGTAATGATGCTACTCTCATTGGATACAGATACTTTATAATCAATAATATATAGAATATT
ATTTATATATAGATAATTATCAGGTGTTATATGCGGTATATCCAACGATAATGGGTCTATATCT
GGCCTAATATCTAATATTATATCTATAAAAGGCACATCATTCCGATACTCTATGTTCAAGGATT
TGCATAACTCACGGCCAAAGTAATCATGTCTAGACATCAGTAGATCTACATCTATGTCCTTTG
CAACACATGCATCCTTAGCTGAATTGATTCTAGCTAGGAATTGTTGGTATTCAGTTTGCTCCAT
CTTAACTGTGTTGTATGATAAAGAATTTGTAAGTTTTATAGGT3’

California encephalitis, 85-415NM virus strain, 523nt; L segment partial polymerase gene:

5’TGGTTAAAATTGATATCGACAACCAATGTCGGGAATAGTTCCTTGAATCTGTCTGAATTTAT
ATGTAATTCCTTACTTACAGGATCTATCCTTACAATAACTATTTCTATAGGGATATTTAATCTT
TGAGATATGTCTCTAGTTAATTCAAAATACTTATCAAATGTTATTAGGCTACTCTCATTAGAAA
CTGAAACCTTGTAATCGATAATATAGAGTATGTTATTTGCAAATAGATAATTATCAGGGGTGA
TATGCGGGAATTCTATGGATAGTGGATCTAGCTCTGGTTTGATGTCTAATATTATATCTACAAA
AGGGACATCATTCCTATATTCTATGTTTAAAGATTTACACAATTCCTTACCAAAGTAGTCGTGC
CTTGCCATAAGCAAATCAACATCTATGTCTTTTGCAACACATGCATCTTTAGCGGAATTAATTC
TTGCAAGGAATTGTTGGTATTCTTGTTGTTCCATAATTGGTATTTTAAATTCTCACTAAAGCCT
TTGTAAGTTTTGTA3’

Fort Sherman, 86MSP18 virus strain, 425nt; L segment partial polymerase gene:

5’CAAATGTATGATCAGTTCTTAAAGAGAATTCAATCAGCTAAAACAGCTACTGTGGCTAAAG
ACATCAGCACAGACATCTTAGAAGCAAGACATGATTATTTTGGCAAAGAACTCTGTGCCTCAA
TTGGAATAGAGTATAAGAACAATGTTCTCTTAGATGAGATTATATTAGATGTTATTCCTGGTGT
CAATCTAATGAACTACAACATACCCAATGTTACTCCAGATAATTATATCTGGGATGGTGATTT
CTTGATAATCCTGGATTATAAAGTGTCAGTTGGGAATGATAGTACTGTACATAACTTACAAGA
AATAGCCTTGATGAGAACACAAGAATTGCTGTCACTGAATTATGATAACATTTGCAAATCAAG
TATAGTGCCTTTTTGTAGGATATTATCATGTAATGATGATGAGGAATT3’

Shokwe, Sa Ar 4042 virus strain, 554nt; L segment partial polymerase gene:

5’AAATTTGTCCAATAACATTTTTCTAAGTTCGAAAAATCTACTGAAATCCAGTTGGATTGGGA
TATTAGGATATCTTTGTTTAAACTCTTCCCCTATAATAGACAACTGATATGTGACTGGATTTGC
CCTTATTATAGCAATTTCTGTATTGATTCCAAGTTCTGACATGACAGGCAGTATTAGACTTGAG
TATTTTTTGTAGGTTATGTCACTGCTATCATTGCCTACAGAAACCTTGTAATCCAGTATAATAA
GAAAATGACCATCCCAGATGTAATTATCTGGTGTCACATGTGGTATACTGTATTTTAGCAGGT
TTACACCTGGTAAAATATCCAGAATTATTTCGTCCAAGAGTACATTGTTTTTGTACTCAATGCC
TAATGAATTGCATAGCTCTCTCCCAAAGTAATCGTGTCTTGCCTCAAGGATATCAGAACTAAT
ATCCTTCGCGACCGTTGCTGTACGCGCAGCTTGGATCCTGTGTAAGAATTGATCATAAGTTTG
GTCCTCCATGTCTGAGATGATTGTCTAATTTAATTATCTTTATGTAG3’

Wyeomyia, Original virus strain, 270nt; L segment partial polymerase gene:

5’TGATGTTGACCTTATGATAGAAAGACATAATTACTTTGGACGGCAATTGTGTGAGTCTCTGG
AAATTGAATATAGGAATGATGTCCCTTTAGTTGACATTTTACTAGATGCAATTCCTGGATTTGA
TCCTATGGTAACTGAAATACCAAATATAACACCTGATAATTACTTGATTATCAATGGCTGGAT
CATAATTATTGATTATAAAGTTTCTGTTAGCAATGAAACAACAGAGATAACACTGGAAAAATA
TAATAAATGTATGGAGAT3’

Xingu, BeH 388464 virus strain, 407 nt; L segment partial polymerase gene:

5’GCACGTCATGATTATTTTGGTCGAGAACTCTGTGCTTCAATAGGCATAGAGTTTAGAAACAA
CGTTCTCTTAGATGAGATCATCCTGGATGTTGCACCTGGAATTAACCTGATGAACTTTAATATA
CCAAATGTTACACCTGATAATTACATCTGGGATGATGATTTCCTAATTATAATGGATTACAAA
GTCTCAGTAGGGAATGATAGTACAGAAATAACATACAAAAAATATACAACCCTGATTCTACC
AGTTATGGAGGAGATAGGTATACCAACTGAAATAGCAATCATAAGGGCAAATCCTGTGACAT
ATCAGATCAGGATAATAGGTGAAAATTTTAAAGCTAAATATCCAGATATACCCATACAGCTAG
ATTTTTCAAAGTTTTTTGATTTGA
GGAAATA3’

Birao, DakArB 2198 virus strain, 389 nt; L segment partial polymerase gene:

5’TACTGGGTTAGCCCTGATAATTGCAATCTCGGTTTCAACACCAATCTCTCCCATGACAGGAA
GGATTAATGAGGTGTATTTCTTATATGTTACTTCTGTGCTATCAGTTCCCACTGATACTTTGTA
GTCCAAGATGATCAAGAACTGACCATCCCAGATGAAATTATCCGGAGTAACATTTGGTATATT
GTATTTAAGCAAGTTTATTCCAGGGGCGATCTCTAATATTATCTCGTCGAGCAATACATTGTTC
CTGAATTCGATATTTAGGGCATTGCAGAGTTCTCTTCCAAAATAATCATGCCGAGCTTCAAGA
ATGTCAGAACTAATGTCTTTAGCAACCGTTGCTGTTCTAGCTGCTTGGATTCTCTTGGAAAATT
GATCATACACAGAACTAATGTCTTTAGCAACCGTTGCTGTTCTAGCTGCTTGGATTCTCTTGGA
AAATTGATCATACA3’

Bozo, DakArB 7343 virus strain, 512 nt; L segment partial polymerase gene:

5’CTATTGAAATCCAATTGTATTGGGATGGTAGGGTAGAGTCTTTTGAATTCATCGCCTATAAT
AGAGATTTGATATGTCACTGGATTTGCTCTTATTATAGCTATCTCAGCATCAATTCCCATCTCA
TTCAAAACAGGCATGATCAAGGTGTTATACTTTTTGAATGTAATATCGCTACTGTCATGGCTTA
CTGATACTTTATAATCAAGGATTATCAAGAATTGGTTCTCCCAAATATAATTGTCTGGAGTCAC
ATTGGGGATATCATAGTTTACTAAATTTACTCCAGGCTTGACATCCAAGATTATATCATCAAG
CAATACATTATTCTTATATTCAATACCCAAAGCATTGCATAACTCTTTACCAAAATAGTCATGC
CTAGCTTCTAATATGTCAGAACTAATATCCTTGGCAACAGTTGCTGTCCTTGCAGCTTGTATCC
TGTTGAGGAACTGGTCATACATGGAGTCTTCCATGGTTTTTGTTTAATTTTTGTAATTTTCTTTT
GT3’

Lokern, FMS 4332 virus strain, 446nt; L segment partial polymerase gene:

5’GGAGGACAATTATATGAACAATATATTAAAAGGATTCAATCAGCAAAAACAGCAACTGTTG
CAAAAGACATCAACAGTGACATTTTGGAAGCACGTCATGACTACTTTGGTCGTGAATTATGCA
TGTCTATTGGGATTGAATATAAGAATAATGTGTTAC
TCGATGAAATAATCTTGGATATTATCCCAGGTGTGAATTTGTTAAATTACAACATACCAAATG
TTACACCAGATAACTATATTTGGGATGGCGAGTATTTGATTATCTTGGATTATAAAGTTTCTGT
AGGACATGATAGTACTGACATAACATATAAAAAATATACTTCTCTTATTCTCCCCGTTATGGA
TCAGCTAGGGATTCCAACAGAAATAGCAATCATTAGAGCTAACCCTGTGACTTACCAAATCAA
CATTATTGGAGAAACCTTcAAAGCTAGATACCC3’

Playas, 75V3066 virus strain, 508nt; L segment partial polymerase gene:

CAAACTTGTCCAGTAATAATTTCCTCAACTCAAAAAACTTTGAAAAATCTAGCTGTATGGGTA
TATTTGGATATTTAGCTTTAAAATTTTCACCTATTATACTGATCTGATATGTCACAGGATTTGC
CCTTATGATTGCTATTTCAGTTGGTATACCTATCTGCTCCATAACTGGTAGAATCAGGGTTGTA
TATTTTTTGTATGTTATTTCTGTACTATCATTCCCTACTGAGACTTTGTAGTCCATTATAATTAG
GAAATCATCATCCCAGATGTAATTATCAGGTGTAACATTTGGTATATTAAAGTTCATCAGGTT
AATTCCAGGTGCAACATCCAGGATGATCTCATCTAAGAGAACGTTGTTTCTAAACTCTATGCC
TATTGAAGCACAAAGTTCTCGACCAAAATAATCATGACGTGCTTCCAAGATGTCAGTGCTGAT
GTCCTTTGCTACAGTGGCTGTTTTAGCTGATTGAATCCTCTTGAGGAATTGATCATACATTTG

Iaco, BeAr 314206 virus strain, 508nt; L segment partial polymerase gene:

5’ATCAAAGATATAATCAGTACCAGGCCAGGATTAATGCTGCTAGGGATCCAACAGTAGCAAA
AGACATTGATGTTGATCTTATGATAGATCGACACAACTATTTTGGCCGTCAGCTATGTGAGGC
TATTAATATACAATATCGAAATGATGTCCCATTATTAGACATATTATTGGATGCTATACCTGGG
CTTGACCCGATGGCATTGGAAATACCAAATATAACACCTGACAATTACATCATATTAGATGGT
CGAATCATCATTATAGATTACAAAGTTTCAGTTAGCTCTGAGACTACGGAAATAACTCTTGAA
AAATATAATAGATGTATGGAGACTATAAAAGAGCAGCTTCCTATAAATTATGAAGTGGCTATT
ATCCGGGTGAACCCAGTAAGCAACCAACTGTTTCTAATAGGTGAGACCTTTATTAGGAATTTC
CCAAATATCCCATTGAACTTAGATTTCTCTCGCTTTTTTGAGCTGAAAACTATGCTCTATCAAA
AATT3’

Santa Rosa, M2-1493 virus strain, 512nt; L segment partial polymerase gene:

5’TGGAGGACAATTATATGAACAATATATTAAAAGGATTCAATCAGCAAAAACAGCAATTGTT
GCAAAAGACATCAACAGTGACATTTTGGAAGCACGTCATGACTACTTTGGTCGTGAATTATGC
ATGTCTATTGGGATTGAATATAAGAATAATGTGTTACTCGATGAAATAATCTTGGATATTATC
CCAGGTGTGAATTTGTTAAATTACAACATACCAAATGTTACACCAGATAACTATATATGGGAT
GGCGAGTATTTGATTATCTTGGATTATAAAGTTTCTGTAGGACATGATAGTACTGACATAACA
TATAAAAAATATACTTCTCTTATCCTCCCTGTTATGGATCAGCTAGGGATTCCAACAGAAATA
GCAATCATTCGAGCTAATCCTGTGACTTACCAAATCAACATTATTGGAGATACCTTTAAAGCT
AGATACCCCAATATACCTATACAATTGGACTTTTCAAAATTTTTTGAATTGAGGAAACTACTCT
TAGATAAGT3’

Alenquer, BeH 301101 virus strain, 354nt; L segment partial polymerase gene:

5’GAGAAGAAAGAATAATTCAATCTATAGTTGAGGCAATATCTAGGTCAATTGGTAGATTCTTT
CCCTCAGATACTCTGTGCAATCCTGGCAACAAGATTAAAATACCTGAAACACACGGAATAAG
AGCTAGAAAACACTGTAAAGGTGCTGTGTGGACTTGTTCCACATCTGATGACGCTAGGAAATG
GAATCAAGGTCACTTTGTTACAAAATTTGCTTTAATGCTGTGCGAGTTCACCCTTCCTAAATGG
TGGCCAATCATCATTAGGGGATGCTCGATGTTCACCAACAAATACATGATGATGAACATGAGG
TATATCAGCATTTTAAGCAGGCACAAGGAATTGGATATTGAAGATGAGTTTTCTAGAACTATA
TTCAAAGCTTTTCATGGTGAGGTGACAGTGCCATGGA3’

Chagres, JW 10 virus strain, 487nt; L segment partial polymerase gene:

5’GCATGGTGGTTTAAGAGAGATCTATGTGATGGGAGCAGAAGAAAGGATAGTCCAAAGTCTA
GTTGAGTCCATCGCAAAGAGTATAGGAAGATTCTTCCCATCTGATACATTGTGCAACCCTGCT
AATAAGACAAAGATCCCTGAAAGCCATGGAGTGCGA
GCCAGGAAGCACTGTGAGGGATCTGTCTGGACATGCTCAACATCTGATGATGCAAGGAAGTG
GAATCAGGGGCACTTTGTGACTAAATTCGCATTGATGCTCAGAGAATTCACACATCCCAAGTG
GTGGCCCATTATTATAAGGGGCTGTTCAATGTTCA
CTGAGAAAAAGATGATGATGAACCTGAACTTTATCAGGATATTGGACTGTCACAAGGAGCTC
AAGACAAGTGATGAGTTTGCAAACACCCTGTTTAAGGCATATCATGGTGAGATAGAGGTCCC
ATGGGCGAAACCTGGAAGGACATATTTGACAACAAAGACAGGG3’

Catu, BeH 151 virus strain, 191nt; L segment partial polymerase gene:

5’TTTTGAAGCTAAAATGTGTATGTATTGTGTGGAAAGAATAGCTAAAGAAAGATGTAAATTA
AACCCAGAGGAGATGATTTCGGAACCTGGCGATGGGAAACTGAAGAAATTGGAAATAAATTC
AGAGAGTGAATTAAGATATTTAATAGAGATGACAAGGAGAGAAATGTCTAAAGAAGAAGAA
TTTATAG3’

Guama, BeAn 277 virus strain, 180nt; L segment partial polymerase gene:

5’TGAAGCTAAAATGTGTATGTACTGTGTTGAAAGAATTGCAAAGGAACGGTGCAAATTGAAT
CCAGAGGAAATGATATCAGAACCAGGAGATGGAAAACTAAAGAAATTAGAAATTAATTCAGA
AAGTGAACTTAGGTATCTCATTGAGATGACGAGACGAGAAATGTCAAAGGAAGAAGA3’

Bwamba, M459 virus strain, 288nt; L segment partial polymerase gene:

5’TTGAAGCTAAAATGTGTATGTATGCTATTGAAAGAATTGCAAAAGAAAGGTGTAGATTAAA
TCCAGAAGAGATGATCTCAGAGCCTGGTGATGGAAAGCTTAGAGTCTTAGAACAAAAAAGCG
AGCAAGAAATTAGATATTTAGTTGAATCAACTAGGCAAAAAAATAGAGATATTGATGATGAA
ATTCAAAATTTAATAGTAAATCCTGAGGAAAACTTAGAAAAAATTGAATTTCTACAAAAGCA
GAAATATAGAGGTCTTAAATTGGAAATAAATGCAGACATGA3’

Pongola, Sa Ar 1 virus strain, 291nt; L segment partial polymerase gene:

5’TTGAAGCTAAAATGTGTATGTATGCAGTGGAGAGGATTGCAAAAGAAAGGTGTAGACTGAA
TCCAGAAGAAATGATATCAGAACCAGGTGACGGGAAATTACGAGTTCTCGAGCAGAAAAGTG
AACAAGAAATTAGATATCTGGTTGAACATACACGTCA
ACGAAATAGAGAAATTGAAGAAGAAATTATGAATTTAGTAGAAGATAATATTGACAAAAATA
TAGATAAATTAGAGATACTACAGACAAGCAGGCATAGAGGGTTGAAAATGGAAATAAATGCA
GACATGGGA3’

Appendix 2, Nucleotide sequences determined in support of international diagnostic and epidemiological efforts presented in Chapter 4. Viral complementary strand sequences are presented in the 5'-3' direction. For full length S segment sequences, the N and NSs ORFs of presented sequences are underlined and highlighted, respectively. For full length M and L segment sequences, the polyprotein and polymerase ORFs appear underlined. GenBank accession numbers appear in parentheses next to the strain designation for each virus sequence.

La Crosse virus isolate Fort Bend County/TX/2009 segment S, 983bp (GU91167)

5'agtagtgtac tccacttgaa tactttgaag ataatttgtt gttgactgtt ttttgcctaa
gggaagttat ctagagtgtg atgtcggatt tggtgtttta tgatgtcgca tcaacaggtg
caaatggatt tgatcctgat gcagggtata tggacttctg tgttaaaaat gcagaatcac
tcaaccttgc tgcagttagg atcttcttcc tcaatgccgc aaaggccaag gctgctctct
cgcgtaagcc agagaggaag gctaacccta aatttggaga gtggcaggtg gaggttatca
ataatcattt tcctggaaac aggaacaatc caattggtaa caacgatctt accatccaca
gattatctgg gtatttagcc agatgggtcc ttgaccagta taatgagaat gatgacgaat
ctcagcgtga attgatcaga acaactatta tcaatccaat tgctgagtcc aacggtgtga
ggtgggacag tgggccagag atctatctat cgtttttccc aggaacagag atgttttgg
aaacctttaa attctacccg ctaaccatcg gcattcacag agtcaagcag ggcatgatgg
accctcaata tctgaaaaag gccttaaggc aacgctatgg caccctcaca gcagacaaat
ggatgtcaca gaaggttgct gcaattgcta agagcctgaa ggatgtagag cagcttaaat
ggggaaaagg aggcctgagc gatactgcta aaacattctt gcagaaattt ggcatcaggc
ttccataaat aaagtatggg gcattcaaat tgggttctaa attctaaatt tctatatgtt
aatttgatca attggttatc caaaagggtt ttcttaaggg aacccacaaa aatagcagct
aaatgggtgg gtggtagggg acagcaaaaa actatagatc aggtcataaa taaaataaaa
tgtattcagt ggagcacact act3'

La Crosse virus isolate Fort Bend County/TX/2009 segment M, 5424bp (GU91168)

5′agtagtgtac taccaagtat agataacgtt taaatattaa agttttggat caaagccaaa
gatgattcgc atgctggtgc tgattgtagt tacggctgca agcccagtgt atcagagatg
tttccaagat ggggctatag tgaagcaaaa cccatccaaa gaggcagtca cagaggtgtg
cctaaaagat gatgttagca tgatcaaaac agaggccagg tatgtaaaaa atgcaacagg
agttttctca aacaatgttg caataaggaa gtggttagtc tctgattggc atgactgcag
acctaagaag atcgttgggg ggcacatcaa tgtaatagaa gttggtgatg acctgtcgct
ccataccgaa tcatatgttt gtagtgcaga ttgtaccata ggtgtagaca aggaaacagc
acaggtcaga cttcaaacag ataccacaaa ccattttgaa atagcaggta ctactgtgaa
gtcaggatgg ttcaaaagca cgacatatat aactctagat cagacctgtg aacacctcaa
agtttcctgc ggcccaaaat cagtacagtt ccatgcctgc ttcaatcagc atatgtcctg
tgtcagattc ttacacagga caatattacc tggttccata gctaattcta tatgtcagaa
cattgaaatc ataattttag ttacacttac cctattaatc ttcatactgt taagcatttt
aagtaggact tatatatgtt atttgttaat gcccatattc atccctatag catatatgta
cggtgtaatt tacaataagt catgcaaaaa gtgcaagcta tgtggcctag tgtatcatcc
attcacagag tgtggtacac attgtgtctg tggtgctcgc tatgatactt ccgatagaat
gaaactgcac agagcttctg gattgtgccc tggttataag agcctaagag ctgccagagt
catgtgcaaa tcaaaagggc ctgcatcaat attgtctata attactgcag tgctggttct
aacctttgtg acaccaatca actccatggt tttaggggaa agtaaggaaa cctttgaact
tgaagagctt ccagacgaca tgctggaaat ggcattaagg ataaattctt attatttcac
ctgtatctta aattatgctg taagctgggg tcttattatt gctggattat tggttggact
gatatttaaa aagtaccagc ataggttctt aaatatttac gcaatgtact gtgaagaatg
tgacatgtat catgacaaat ctgggctgaa gagacatggt gacttcacca acaaatgtag
acaatgcaca tgtggtcaat atgaagatgc tgcaggtttg atggctcata ggaaaaccta
taactgctta gtgcagtaca aagcaaagtg gatgatgaat ttcctaataa tttacatatt
cttaattttg atcaaagatt ctgccatagt tgtacaagct actgggacag atttcaccac
ttgcctagag actgagagta taaattggaa ctgcactggg ccattcctaa accttggaaa
ctgccaaaag caacaaaaga aagagcctta cacaaatatt gcaacccaac taaaagggct
aaaagcaatt tccgtactag atatccctat aattacgggt atcccagacg acattgcagg
cgctttaaga tatatagagg agaaggaaga tttccatatc caactaacta cagaatatgc
aatgctaagt aagtactgtg actattatac ccagttctca gataactcag gatatagtca
aacaacatgg agagtatact tgagatctca tgattttgag gcatgcatac tatatccaaa
tcagcatttc tgcaggtgtg tgaaaaatgg tgagaagtgt agcagctcca attgggactt
tgccaatgaa atgaagaatt attactctgg aaaacaagca aaatttgata aagatttaaa
cctagctcta acagctttgc atcatgcctt cagagggacc tcatctgcat atatagcaac
aatgctctca aagaagtcca atgatgattt gatcgcatat acaaataaaa taaaagcaaa
attcccaggt aatgcattat tgaaggctat aatagattat atagcataca tgaaaggttt
gccagatatg gcaaatttca aatatgatga gttctgggat gaattattgt acaaacctaa
cccagcaaaa gcctcaaatc ttgctagagg aaaagaatca tcctacaact tcaagttagc

aatttcgtca aaatctataa aaacctgtaa gaatgttaag gatgttgcct gcttatcacc
aaggtcaggt gctatatact cttcaataat tgcatgtggt gaacccaacg ggccaagtgt
gtatagaaaa ccatcaggtg gtgtatttca atccagtact gaccggtcta tatattgttt
gctagacagt cattgtttag aggagtttga ggctatcagt caagaggagt tggatgcagt
aaaaaaatcc aagtgttggg agattgaata tcctgacgta agacccctcc aagaaagtga
cggggctaaa agctgtagaa tgaaagattc tgggaactgt aatgttgcaa ctaacagatg
gccagtaatg caatgcgaga atgacaaatt ttattactca gaacttcaaa aagattatga
caaaactcaa gatattggtc attattgctt aagccctgga tgtaccacta tccggtaccc
tattaatcca aagcatatct ctaattgtaa ttggcaggta agcagatcca gcatagcaaa
gatagatgta cataatgttg aagatattga acaatataaa aaggctataa cccaaaagct
tcaaacaagc ctatctctat tcaagtatgc aaaaacaaaa aacctgccac acatcaagcc
aatttataaa tacataacta tagaaggaac agaaactgca gaagggatag agagcgcata
tattgaatca gaggtgcctg cattggctgg gacatctatt gggttcaaaa ttaattctaa
agagggcaaa cacttgctag atgttatagc atacgtaaaa agtgcttcgt actcttcagt
atatgcgaaa ttatactcaa ctggcccaac atcaggaata aatactaaac atgatgaatt
gtgtactggc ccatgcccag caaatatcaa tcatcaagtt gggtggctaa catttgcgag
agagaggaca agttcatggg gatgcgaaga gtttggttgc ctggctgtaa gtgatgggtg
cgtgttcggg tcatgccaag acataataaa agaagaacta tctgtctata gaaaagagac
cgaagaagtg accaatgtgg agctatgtct aacattttca gacaagacat attgcacaag
cttaaaccct gttaccccta ttataacaga tctatttgag gtacagttta aaactgtaga
gacttacagc ttgcccagaa ttgttgctgt gcaaaatcat gagattaaga ttgggcaaat
aaatgatcta ggggtttact ctaagggctg tgggaatgtt caaaaggtta atggaactgt
ttatggcaat ggggtcccca gatttgatta cttatgtcat ttagctagta gaaaagaagt
catcgttaga aaatgctttg ataatgatta ccaagcatgc aaattccttc aaagccctgc
tagttataga cttgaagaag acagcggtac tgtaactata attgattata agaagatatt
aggcacaatc aagatgaagg caattttagg agatgtcaaa tacaaaacat ttgctgacag
tgttgatata actgcagaag ggtcgtgtgc cggctgtatc aattgctttg agaatattca
ttgtgaatta acgctgcata ccacaattga agccagctgc ccaattaaaa gttcttgcac
agtattccat gataggattc ttgtaacccc aaatgaacac aaatatgcat taaaaatggt
atgcacagaa aagccaggga atacactcac cattaaagtc tgcaatacta gaattgaagc
ttcaatggcc cttgttgatg caaagcctat catagaacta gcaccagttg atcagacggc
atatataaga gaaaaagatg aaagatgcaa aacttggatg tgtagggtga gggatgaagg
gctgcaggtt atcttagagc catttaaaaa cttgtttgga tcttatattg ggatatttta
cacatttatt atatctataa tagcattatt ggttatcatc tatgtactac tgcctatatg
cttcaaatta agggataccc ttagaaagca cgatgatgca tataagagag agatgaaaat
tagataaggg attcatgtaa aacaaaattg agtcctgtat tatatattct atttgtagta
tagctgttgt taagtggggg gtggggaact aacaacagcg taagtttatt ttgcaaacat
tatttatacg tggagcacac tact3’

La Crosse virus isolate Fort Bend County/TX/2009 segment L, 6979bp (GU91169)

5′agtagtgtac tcctatctac aaaacttaca aaaaattcag ttatatcaca atatacacat
aatggactat caagagtatc aacaattctt ggctagaatc aacactgcaa gagacgcatg
tgtagccaaa gatatcgacg ttgacctatt aatggccaga cacgattatt ttggaagaga
gctatgcaag tccttaaata tagaatatag aaatgatgta ccatttgtag acataatttt
agacataaag cctgaaatag acccactaac catagacgca ccatatatta ccccagataa
ttacttgtat ataaataatg tattgtacat catagattat aaggtctctg tatcgaatga
aagtagtttg ataacatatg acaaatacta cgagttaact agagatatat ctgacagatt
aggtattccg atagaaatag tcatcatccg tatagaccct gtgagcaaag atttacatat
caattctgat aggtttaaag aactctaccc tgcaatagtg gtggatataa acttcaatca
gttttttgat ttgaagcaat tgctctatga aaaattcggt gatgatgaag aatttttact
aaaagtagcg catggtgatt tcactctcac agcaccctgg tgcaagactg ggtgccctga
tttttggaag catcctattt ataaagagtt caagatgagt atgccagtac ctgaacgaag
gctttttgaa gaatcagtca aattcaatgc ttacgaatca gagagatgga ataccaactt
agttaaaatt agagaatata caaagaaaga ctatttagaa cacatttcaa aatctgcaaa
gaatattttt ctggctagtg gattctataa acagccaaat aaaaacgaga ttagcgaggg
atggacatta atggtcgaaa gggtgcaaga ccagagggaa ttctcaaaaa ccctccatga
tcagaaaccc agcatacatt ttatatgggg ggcccataac ccaggaaaca gtaacaatgc
aactttcaaa ctcatactgc tctcaaagtc cctacaaagc ataaagggta tatctaccta
cacagaagca tttaaatctt taggtaaaat gatggatatc ggggataaga ctattgaata
tgaagaattt tgcatgtccc taaaaagcaa agcaagatca tcatggaagc aaataatgaa
taaaaagtta gagcccaaac agataaataa tgcacttgtc ctatgggaac agcaatttat
ggtaaataat gacctaatag acaaaagcga aaaattgaaa ttgtttaaga acttctgcgg
tataggcaag cacaagcaat tcaagaataa gatgctagaa gacatagaga tgtcaaagcc
caagatatta gactttgatg acccaaatat gtatctggct agtttaacca tgatggaaca
aagtaagaag atattatcca aaagtaatgg gttgaaacca gataatttta tattgaatga
atttggatcc agaattaaag atgctaataa ggaaacatat gataacatgc ataaaatttt
tgagacaagg tattggcaat gcatatctga cttctccacc ctaatgaaga atatcttatc
tgtgtctcaa tataataggc acaatacatt taggatagct atgtgtgcta ataacaatgt
gtttgctata gtatttcctt cagctgatat aaaaacaaaa aaagcaactg tagtttatag
tatcatagta ctgcacaaag aggaggaaaa catattcaac ccaggatgtt tgcatggcac
attcaagtgc atgaatggat acatttccat ttctagagct ataaggctag acaaagagag
gtgccaaaga attgtttcct cacccgggct atttttaact acttgcttac tattcaaata
tgacaatcca accctagtga tgagtgatat catgaatttt tctatataca ctagcttgtc
tatcacaaag agtgttttat ctttgacaga gccagcacga tacatgatta tgaattcatt
agccatctcc agcaatgtta aagattatat agcagagaag ttttcccctt atacaaagac
actatttagc gtctatatga ccagattaat aaaaaatgct tgctttgatg cttatgacca
gagacaacgt gttcaactta gagatatata tttatccgat tatgacataa cccaaaaagg
tattaaagac aatagagaat taacaagtat atggttccct ggtagtgtaa cactaaaaga
atatttaaca caaatatatt taccatttta tttcaatgct aaaggattac atgagaaaca

tcatgttatg gtagatctag caaagactat attagaaata gaatgcgaac agagggaaaa
cataaaggaa atatggtcta caaattgtac caaacaaaca gtgaatctta aaattctaat
ccattccttg tgcaagaatc tacttgcaga cacctcaaga cacaatcatt tgcggaatag
gatagagaat aggaacaatt tcagaaggtc tataacaact atttcaacat tcacaagttc
aaagtcttgt ctcaaaatag gagactttaa gaaagagaaa gaattacaat cagtaaaaca
aaagaaaatc ttagagacac agagtcgtaa aatgaggtta gcaaatccaa tgtttgtgac
agatgagcaa gtatgccttg aagtcagtca ctgcaattat gagatgctaa gaaatgctat
gccaaattat acagattaca tatcaactaa agtattcgat agattatatg agttgttgga
taaaggagtt ttaacagata aacccgttat agagcaaata atggatatga tggttaacca
caagaagttc tacttcacat ttttcaataa gggccagaaa acatcaaaag atagagagat
atttgttgga gaatatgaag ctaaaatgtg catgtatgca gttgaaagga tagcaaaaga
aagatgcaaa ctaaaccctg atgaaatgat atctgagcca ggagatggta agttgaaggt
gttagaacaa aaatcagaac aagaaattag attcttggtt gagactacaa gacaaaaaaa
ccgtgaaata gatgaagcaa ttgagacgtt agctgcagaa ggatatgaga gtaatctaaa
gaagattgaa aagctttcac ttggcaaagc aaagggccta aagatggaaa taaatgcaga
catgtctaaa tggagtgccc aagatgtttt ttacaaatac ttttggctca tagccttgga
ccctatcctc tacccacaag aaaaagagag aatattgtac tttatgtgca attacatgga
taaggagttg atactgccag atgaattatt attcaattta ttagaccaaa aagttgcata
ccaaaatgac ataatagcta ctatgactaa tcaattaaat tccaacacgg tctcaataaa
gaggaattgg cttcaaggaa attttaacta tacctcaagt tacgtccata gctgcgcaat
gtctgtgtat aaagaaatat taaaagaggc tataacatta ctggacgggt ctatattggt
caattcatta gttcattcag atgacaacca aacatcaata acaatagttc aggacaaggt
ggaaaatgat aaaattatag actttgcaat gaaagaattt gagaaagctt gtttgacatt
tggatgccaa gcaaatatga aaaagacata tgttacaaac tgtattaaag agtttgtttc
attgtttaat ttgtatggtg agcccttctc aatttatggc agattcttat taacatctgt
gggcgattgt gcttatatag ggccttacga ggatttagct agtcgaatat catcagccca
gacagccata aaacatggtt gcccgcccag tctggcatgg gtatctatag caataagcca
ctggatgact tctctaacat acaacatgct accagggcaa tcaaatgacc cgcttgatta
ttttccagca gagagtagga aagacatacc tatagaattg aatggcgtat tagatgctcc
attatcaatg atcagtacag ttggattaga atctgggaat ttatactttt taataaaatt
gttgagcaaa tacaccccag ttatgcaaaa aagggagtca gtagtcaatc agatagctga
agttaagaac tggaaagttg agaatctaac agataatgaa gtattcagac taaagatact
tagatatttg gttttagatg cagagatgga ccctagtgac attatgggtg agacaagtga
tatgagaggg agatctattt taacacctag aaaattcaca acagtaggta gcttaagaaa
attgtattca tttagtaagt atcaagatag actgtcttcc cctggaggca tgattgaatt
atttacctat ttgcttgaga agcctgaatt gttagtgacc aaaggggaag atatgaaaga
ttatatggaa tctgtgatat tccgatataa ctccaaaagg ttcaaggaaa gtttgtcaat
acagaatcca gcacagctat ttatagaaca gatattgttc tcgcataagc ccataataga
cttttctggc attagggaca aatatataaa cctacatgat agcagagccc tagaaaaaga
acctgatata ctaggaaaag taacgtttac agaggcttat aggttgttaa tgagggatct
atctagctta gaactgacca atgacgatat tcaagtgatt tattcttata taatactcaa

tgatcctatg atgataacaa ttgcaaacac acatatacta tcaatatacg ggagtccaca
aagacggatg ggtatgtcct gttctacgat gccagaattt agaaatttga agttgataca
ccattcccca gccttggttt taagagcata tagtaaaaat aatccagaca tccaaggtgc
cgaccctaca gaaatggcta gagatttagt tcatctaaaa gaatttgttg aaaatacaaa
tttggaagag aaaatgaagg tcagaattgc tataaatgag gctgaaaagg ggcaacggga
tatagtcttt gaactaaaag agatgactag attttatcag gtttgctatg agtatgttaa
gtctacagaa cacaagataa aagtcttcat ccttccagca aagtcataca caacaacaga
tttctgctcg ctcatgcaag ggaacttaat aaaggacaaa gagtggtaca cagtccacta
cttgaagcag atattgtctg gtggtcacaa agccataatg cagcataatg ccactagtga
acaaaatatt gcttttgagt gttttaagtt gataacccat tttgcagact cattcataga
ttcgctatct agatcagcct ttttacagtt gataatagat gaatttagtt ataaagatgt
gaaggttagt aagctctatg acataataaa gaatgggtat aatcgaactg acttcatacc
attgcttttt agaaccggcg atttaagaca agcagactta gataaatatg atgctatgaa
aagccatgag agagtaacat ggaatgactg gcagacatct cgtcatttgg acatgggctc
tattaatcta acaataactg gttataatag gtcaataaca ataattggag aagacaacaa
gttgacatat gcagagctat gtctgaccag gaaaactcct gagaatataa ctataagcgg
tagaaagttg ctaggctcaa ggcacggtct taaatttgaa aatatgtcta aaatccaaac
atacccaggc aactattaca taacatatag gaagaaggat cgccatcagt ttgtatacca
gatacattct catgaatcaa taacaagaag gaacgaagaa cacatggcta tcaagactag
aatatacaat gaaataaccc cggtatgtgt agttaatgtt gcagaggtgg atggggatca
acgtatatta ataagatcct tagattattt aaataatgat atattctctc tttcaagaat
tagagttggg cttgatgaat ttgcaacaat aaaaaaagca cacttcagca agatggtttc
atttgaagga cccccaatta agacaggcct tcttgatctt actgaattga tgaaatccca
agatttactt aaccttaatt atgacaatat aaggaatagc aacctgatat ccttttcaaa
actaatttgc tgtgaagggt cagataatat aagtgatggc ctagagttcc tgtctgatga
tcctatgaac tttacagaag gtgaagcaat acattcaaca ccaattttta atatatatta
ctcaaagaag ggagataaac atatgacata cagaaatgca attaaactac tgatagagag
ggaaactaag atttttgaag aagcctttac attcagcgag aatggcttca tatcaccgga
gaatcttggt tgcttagaag ctgtggtatc attaataaaa ttgttgaaaa ctaatgaatg
gtccacagtt atagacaaat gtatccatat atgcttaata aaaaatggca tggaccatat
gtatcattca tttgatgttc ccaaatgttt tatggggaat cctattacta gagatataaa
ttgggtgatg tttagagaat ttatcaatag cttaccagga acagatatac caccatggaa
tgttatgaca gagaacttca aaaagaaatg tattgcccta ataaattcca aatttgaaac
acaaagagat ttctcagaat tcactaaatt gatgagaaag gaaggtggaa ggagcaatat
agagtttgac tagtagttat gggtttatag gggactgaca attagaccat aaatctggga
ggggtttgaa attggctaag attcaaaaag aggggaatta acagcaactg tataaatttg
tagataggag tacactact3’

La Crosse virus isolate Dallas/TX/2009 segment S, 983bp (GU591164)

5′agtagtgtac tccacttgaa tactttgaag ataatttgtt gttgactgtt ttttgcctaa
gggaagttat ctagagtgtg atgtcggatt tggtgtttta tgatgtcgca tcaacaggtg
caaatggatt tgatcctgat gcagggtata tggacttctg tgttaaaaat gcagagtcac
tcaaccttgc tgcagttagg atcttcttcc tcaatgccgc aaaggccaag gctgctctct
cgcgtaagcc agagaggaag gctaacccta aatttggaga gtggcaggtg gaggttgtca
ataatcattt tcctggaaac aggaacaatc caattggtaa caacgatctt accatccaca
gattatctgg gtatttagcc agatgggtcc ttgaccagta taatgagaat gatgacgaat
ctcagcgtga attgatcaga acaactatta tcaatccaat tgctgagtcc aatggtgtga
ggtgggacag tgggccagag atctatctat cgtttttccc aggaacagag atgttttgg
aaacctttaa attctacccg ctaaccatcg gcattcacag agtcaagcag ggcatgatgg
accctcaata tctgaaaaag gccttaaggc aacgctatgg caccctcaca gcagacaaat
ggatgtcaca gaaggttgct gcaattgcta agagcctgaa ggatgtagag cagcttaaat
ggggaaaagg aggcctgagt gatactgcta aaacattcct gcagaaattt ggcatcaggc
ttccataaat aaagtatggg gtattcaaat tgggttctaa attctaaatt tctatatgtt
aatttgatca attggtcatc caaaagggtt ttcttaaggg aacccacaaa aatagcagct
aaatgggtgg gtggtagggg acagcaaaaa actatagatc aggtcataaa taaaataaaa
tgtattcagt ggagcacact act3′

La Crosse virus isolate Dallas/TX/2009 segment M, complete sequence, 4525bp (GU591165)

5′agtagtgtac taccaagtat agataacgtt taaatattaa agttttggat caaagccaaa
gatgattcgc atactagtgc tgattgtagt tacggctgca agcccggtgt atcagagatg
tttccaagat ggggctatag tgaagcaaaa cccatccaaa gaggcagtca cagaggtgtg
cctaaaagat gatgttagca tgatcaaaac agaggccagg tatgtaaaaa atgcaacagg
agttttctca aacaatgttg caataaggaa gtggttagtc tctgattggc atgactgcag
acctaagaag atcgttgggg ggcacatcaa tgtaatagaa gttggtgatg acctgtcgct
ccataccgaa tcatatgttt gcagtgcaga ttgtaccata ggtgtagaca aggaaacagc
acaggtcaga cttcaaacag atactacaaa ccattttgaa atagcaggta ctactgtgaa
gtcaggatgg ttcaaaagca cgacatatat aactctagat cagacctgtg aacacctcaa
agtttcctgc ggcccaaaat ctgtacagtt ccatgcctgc ttcaatcagc atatgtcttg
tgtcagattc ttacacagga caatattgcc tggttccata gctaattcca tatgtcagaa
tattgaaatc ataattttag ttacacttac cttattaatc ttcatactgt taagcatttt
aagtaagact tatatatgtt atttgttaat gcccatattc atccccatgg catatatgta
cggtgtaatt tacaataagt catgcaaaaa gtgcaagcta tgtggcctag tgtatcatcc
attcacagag tgtggcacac attgtgtctg tggtgctcgc tatgatactt ccgatagaat
gaaactgcac agagcttctg gattgtgccc tggttataaa agcctaagag ctgccagagt
catgtgcaaa tcaaaagggc ctgcatcaat attgtctata attactgcag tgctggttct
gacatttgtg acaccaatca actctatggt tttaggggaa agtaaggaaa cctttgaact

tgaagagctt ccagacgaca tgctggaaat ggcattaagg ataaattctt actacttcac
ctgtatctta aattatgctg taagctgggg tcttattatt gctggattat tggttggact
gatatttaaa aagtaccagc ataggttctt aaatatttac gcaatgtact gtgaagaatg
tgacatgtat catgacaaat ctgggctgaa gagacatggt gacttcacca gcaaatgtag
acaatgcaca tgtggtcaat atgaagatgc tacaggtttg ataactcaca ggaaaaccta
taactgctta gtgcagtaca aagcaaagtg gatgatgaat ttcctaataa tttacatatt
cttaattttg atcaaagatt ctgccatagt tgtacaagct actgggacag atttcaccac
ctgcctagag actgagagta taaattggaa ctgcactggg ccattcctaa accttggaaa
ctgccaaaag caacaaaaga aagaacctta cacaaatatt gcaacccagc taaaagggct
aaaagcaatt tccgtactag atatccctat aattacgggt atcccagacg acattgcagg
tgctttaaga tatatagagg agaaggaaga tttccatgtc caactaacta cagaatatgc
aatgctgagt aagtactgtg actattatac ccagttctca gataactcag gatatagtca
aacaacatgg agagtatact tgagatctca tgattttgag gcatgcatac tatatccaaa
tcagcatttc tgcaggtgtg tgaaaaatgg tgagaagtgt agcagctcta attgggactt
tgccaatgaa atgaagaatt attactctgg gaaacaagca aaatttgata aagatttaaa
cctagctcta acagctttgc atcatgcctt cagagggacc tcatctgcat atatagcaac
aatgctctca aagaagtcca atgatgattt ggttgcatat acaaataaga taaaagcaaa
attcccaggt aatgcattat tgaaggctat aatagattat atagcataca tgaaaggttt
gccagatatg gcaaatttca aatatgatga gttctgggat gaattattgt acaaacctaa
cccagcaaaa gcctcaaatc ttgctagagg aaaagagtca tcctacaact tcaagttggc
agtttcgtca aaatctataa gaacctgtaa gaatgttaag gatgttgcct gcttatcacc
aaggtcaggt gctatatact cttcaataat tgcatgcggt gaacccaacg ggccaagtgt
gtatagaaaa ccatcaggtg gtgtatttca atccagtact gaccggtcta tatattgttt
gctagacagc cattgtctag aggagtttga ggctatcagt caagaggagt tggatgcagt
aaaaaaatcc aagtgttggg agattgaata tcctgacata agacccctcc aagaaagtga
cgggtctaaa agctgtagaa tgaaagattc tgggaactgt aatgttgcaa ctaacagatg
gccagtaatg caatgcgaga atgacaaatt ttattactca gaacttcaaa aagattatga
caaaactcaa gatattggtc actattgctt aagccctgga tgtaccacta tccggtaccc
tattaatcca aagcatatct ctaattgtaa ttggcaggta agcagatcca gcatagcaaa
gatagacgta cataatgttg aagatattga acaatataaa aaggctataa cccaaaagct
tcaaacaagc ctatctctat tcaagtatgc aaaaacaaaa aacctgccac acattaagcc
aatttataaa tacataacta tagaaggaac agaaactgca gaagggatag agagcgcata
tattgaatca gaggtgcctg cattggctgg gacatctatt gggttcaaaa ttaattctaa
agagggcaaa cacttgctag atgttatagc atatgtaaaa agtgcatcgt attcttcagt
atatgcgaaa ttatactcaa ctggcccaac atcaggaata aatactaaac atgatgaatt
gtgcactggc ccatgcccag caaatatcaa tcatcaagtt gggtggctaa catttgcgag
agagaggaca agttcatggg gatgcgaaga gtttggttgc ctggctgtaa gtgatgggtg
cgtgttcggg tcatgccaag acataataaa agaagaacta tctgtctata ggaaggatac
tgaagaagtg accaatgtgg agctatgtct aacattttca gacaagacat attgcacaaa
cttaaaccct gttaccccta ttataacaga tctatttgag gtacagttta aaactgtaga
gacttacagc ttgcccagaa ttgttgctgt gcaaaaccat gagattaaga ttgggcaaat

aaacgatcta ggggtttact ctaagggctg tgggaatgtt caaaaggtta atggaactgt
ttatggcaat ggggtcccca gatttgatta cttatgtcat ttagctagta gaaaggaagt
catcgttaga aaatgctttg ataatgatta ccaagcatgc aaattccttc aaagccctgc
tagttataga cttgaagaag acagcggtac tgtaactata attgattata agaagatatt
aggcacaatc aagatgaagg caattttagg agatgtcaaa tacaaaacat ttgctgacag
tgttgatata actgcagaag ggtcgtgtgc tggctgtatc aattgctttg agaatattca
ttgtgaatta acactgcata ccacaattga agccagctgc ccaattaaaa gttcttgcac
agtattccat gataggattc ttgtaacccc aaatgaacac aaatatgcat taaaaatggt
atgcacagaa aagccaggga atacactcac cattaaagtc tgcaatacta gaattgaagt
ttcaatggcc cttgttgatg caaagcctat catagaacta gcaccagttg atcagacggc
atatataaga gaaaaagatg aaagatgcaa aacttggatg tgtagggtga gggatgaagg
gctgcaggtc atcttggagc catttaaaaa tttatttgga tcttatattg gaatatttta
cacatttatt atatctataa tagcattatt ggttatcatc tatgtactac tgcctatatg
cttcaaatta agggataccc ttagaaagca cgatgatgca tataagagag agatgaaaat
tagataaggg attcatgtaa aacaaaattg agtcctgtat tatatattct atttgtagta
tagctgttgt taagtggggg gtggggaact aacaacagcg taagtttatt ttgcaaacgt
tatctatact tggtagtaca ctact3′

La Crosse virus isolate Dallas/TX/2009 segment L, complete sequence, 6979bp (GU591166)

5′agtagtgtac tcctatctac aaaatttaca aaaaattcag ttatatcaca atatacacat
aatggactat caagagtatc aacaattctt ggctaggatc aacactgcaa gagacgcatg
tgtagccaaa gatatcgacg ttgacctatt aatggccaga cacgattatt ttggaagaga
gctatgtaag tccttaaata tagaatatag aaatgatgta ccattcgtag atataatttt
agacataagg cctgaaatag acccactaac catagacgca ccacatatta ccccagataa
ttacttgtat ataaataatg tattgtacat catagattat aaggtctctg tatcgaatga
aagtagtttg ataacatatg acaaatacta cgagttaact agagatatat ctgacagatt
aggtattccg atagaaatag tcatcatccg tatagatcca gtgagcaaag atttacatat
caattctgat aggtttaaag aactctaccc tgcaatagtg gtggatataa atttcaatca
gtttttgat ttaaagcaat tgctctatga aaaattcggt gatgatgaag aatttttact
aaaagtagcg catggtgatt tcactctcac agcaccctgg tgcaagactg ggtgccctga
tttttggaag catcccattt ataaagagtt caagatgagt atgccagtac ctgagcgaag
gctttttgaa gaatcagtca aattcaatgc ttacgaatca gagagatgga atactaactt
ggttaaaatt agagaatata caaagaaaga ctatttagaa cacatttcaa aatctgcaaa
gaatattttt ctggctagtg gattctataa acagccaaat aaaaacgaga ttagtgaggg
atggacatta atggttgaaa gggtgcaaga ccagagggaa atctcaaaaa ccctccatga
tcagaaaccc agcatacatt ttatatgggg ggcccataac ccaggaaaca gtaacaatgc
aaccttcaaa ctcatactgc tctcaaagtc cctacaaagc ataaagggta tatctaccta
cacagaagca tttaaatctt taggtaaaat gatggatatc ggggataaga ctattgaata
tgaagaattt tgcatgtccc taaaaagcaa agcaagatca tcatggaagc aaataatgaa

taaaaaatta gagcccaaac agataaataa tgcacttgtc ctatgggaac agcagtttat
ggtaaacaat gacctaatag acaaaagcga aaaattgaaa ttgtttaaga acttctgtgg
tataggcaag cacaagcaat tcaagaataa gatgctagaa gacattgaga tgtcaaagcc
caagatatta gactttgatg acccaaatat gtatctggct agcttaacca tgatggaaca
aagtaagaag atattgtcta aaagtaatgg gttgaaacca gataatttta tattgaatga
atttggatcc agaattaaag atgctaataa ggaaacatat gataacatgc ataaaatttt
tgagacaaga tattggcaat gcatatctga cttctccacc ctaatgaaga atatcttatc
tgtgtctcaa tataataggc acaatacatt taggatagct atgtgtgcta ataacaatgt
gttcgctata gtatttcctt cagctgatat aaaaacgaaa aaagcaactg tagtttatag
tatcatagtg ctgcacaaag aggaggaaaa catattcaac ccaggatgtt tgcatggcac
attcaagtgt atgaatgggt acatttccat ttctagagct ataaggctag ataaagagag
gtgccaaaga attgtttctt cgcctggatt atttttgact acttgcttac tattcaaaca
tgataatcca accctagtga tgagtgatat catgaatttt tctatataca ctagcttgtc
tatcacaaag agtgttttat ctttgacaga gccagcacga tacatgatta tgaattcatt
agccatctcc agtaatgtta aagattatat agcagagaag ttttcccctt atacaaagac
actatttagc gtctatatga ccagattaat aaaaaatgct tgctttgatg cttatgacca
gagacagcgt gttcaactta gagatatata tttatccgat tatgacataa cccaaaaagg
tattaaagac aatagagaat taacaagcat atggttccct ggtagtgtaa cgctaaaaga
atatttaaca caaatatatt taccatttta tttcaatgct aaaggattac atgagaaaca
tcatgttatg gtagatctag caaagactat attagaaata gaatgcgaac agagggaaaa
cataaaggaa atatggtcta caaattgtac caaacaaaca gtgaatctta aaattctaat
ccattccttg tgcaagaatc tgcttgcaga cacctcaaga cacaatcatt tgcggaacag
gatagagaat aggaataatt tcagaaggtc tataacaact atttcaacat tcacaagttc
aaagtcttgt ctcaaaatag gagactttaa gaaagagaaa gaattacaat cagtaaaaca
aaagaaaatc ttagagacac aaagtcgcaa aatgaggtta gcaaatccaa tgtttgtgac
agatgaacaa gtgtgccttg aagtcagtca ctgcaattat gagatgctaa gaaatgctat
gccaaattat acagattaca tatccactaa agtgttcgat agattatatg aattgttgga
taaaggagtt ttaacagata aacccgttat agagcaaata atggatatga tggttaacca
caagaagttc tacttcacat ttttcaataa gggccagaaa acatcaaaag atagagaaat
atttgttgga gaatatgaag ctaaaatgtg catgtatgca gttgaaagga tagcaaaaga
aagatgcaaa ctaaaccctg atgaaatgat atctgagcca ggagatggta agttgaaggt
gttggaacaa aaatcagaac aagaaattag attcttggtt gagactacaa gacaaaaaaa
ccgtgaaata gatgaagcaa ttgagacgtt agctgcagaa ggatatgaga gtaatctaaa
gaagattgaa aagctttcac ttggcaaagc aaagggccta aagatggaaa taaatgcaga
catgtctaaa tggagtgccc aagatgtttt ttacaaatac ttttggctca ttgccttgga
ccctatcctc tacccacaag aaaaagagag aatattatac tttatgtgca attacatgga
taaggagttg atactgccag atgaattatt attcaattta ttagaccaaa aagttgcata
ccaaaatgac ataatagcta ctatgactaa tcaattaaat tccaacacgg ttttaataaa
gaggaattgg cttcaaggaa attttaacta tacctcaagt tacgtccata gctgcgcaat
gtctgtgtat aaagaaatat taaaagaggc tataacatta ctggacgggt ctatattggt
caattcatta gttcattcag atgataacca aacatcaata acaatagttc aggacaaggt

ggaaaatgat aaaattatag actttgcaat gaaagaattt gagaaggctt gtttgacatt
tggatgccaa gcaaatatga aaaagacata tgttacaaac tgtattaaag agtttgtttc
attgtttaat ttgtatggtg agccctttc aatttatggc agattcttat taacatctgt
gggcgattgt gcttatatag ggccttacga agatttagct agtcgaatat catcagccca
gacagccata aaacatggtt gcccgcccag tctggcatgg gtatctatag caataagcca
ctggatgact tctctaacat acaatatgct accagggcaa tcaaatgacc cgcttgatta
ttttccagca gagagtagga aagacatacc tatagaattg aatggcgtat tagatgctcc
attatcaatg atcagtacag ttggattaga atctgggaat ttatactttt taataaaatt
gttgagcaaa tacaccccag ttatgcaaaa aagggagtca gtagtcaatc agatagctga
agttaagaac tggaaagttg agaatctaac agataatgaa gtattcagac taaagatact
tagatatttg gttttagatg cagagatgga ccctagtgac attatgggtg agacaagtga
tatgagaggg agatctattt taacacctag aaaattcaca acagcaggta gcttgagaaa
attgtattca tttagtaagt atcaagatag actgtcttcc cctggaggca tgattgaatt
atttacctat ttgcttgaga agcctgaatt gttagtgacc aaaggggaag atatgaaaga
ttatatggaa tctgtgatat tccgatataa ctccaaaagg ttcaaggaaa gtttatcaat
acagaatcca gcacagctgt ttatagaaca gatattgttc tcacataagc ccataataga
cttttctggc attagggaca aatatataaa cctacatgat agcagagccc tagagaaaga
gcctgatata ttagggaaag taacgtttac agaggcttat agattgttaa tgagggatct
atctagctta gaactgacaa atgacgatat tcaagtaatt tattcttata taatactcaa
tgatcctatg atgataacaa ttgcaaacac acatatacta tcaatatacg ggagtccaca
aagacggatg ggtatgtcct gttctacaat gccagaattt agaaatttaa agttgataca
ccattcccca gccttggttt taagagcata cagtaaaaat aatccagaca tccaaggtgc
cgaccccaca gaaatggcta gagatttagt tcatctaaaa gaatttgttg aaaatacaaa
tttggaagag aaaatgaagg tcagaattgc tataaatgag gctgaaaagg ggcaacggga
tatagtcttt gaactaaaag agatgactag attttatcag gtttgctatg agtatgttaa
gtctacagaa cacaagataa aagtcttcat ccttccagca aagtcataca caacaacaga
tttctgctcg ctcatgcaag ggaacttaat aaaggacaaa gagtggtaca cagtccacta
cttgaagcag atattgtctg gtggtcacaa agccataatg cagcataatg ccactagtga
gcaaaatatt gcttttgagt gttttaagtt gataacccat tttgcagact cattcataga
ttcgctatct agatcagcct ttttacagtt gataatagat gaatttagtt ataaagatgt
gaaggttagt aagctctatg acataataaa gaatgggtat aatcgaactg actttatacc
tttgcttttt aggaccggcg atttaagaca agcagactta gataaatatg atgctatgaa
aagccatgag agagtaacgt ggaatgactg gcagacatct cgtcatttgg acatgggctc
tattaatcta acaataactg gttataatag gtcaataaca ataattggag aagacaataa
gttgacatat gcagagctat gtctgaccag gaaaactcct gagaatataa ctataagcgg
tagaaagttg ctgggctcaa ggcacggtct taaatttgaa aatatgtcta aaatccaaac
atacccaggc aattattaca taacatatag gaagaaggat cgccaccagt ttgtatacca
gatacattct catgaatcaa taacaaggag gaacgaagaa cacatggcta tcaagactag
aatatacaat gaaataaccc cggtatgtgt agttaatgtt gcagaggtgg atggggatca
acgtatatta ataagatcct tagattattt aaataatgat atattctctc tttcaagaat
taaagttggg cttgatgaat ttgcaacaat aaaaaaagca cacttcagca agatggtttc

atttgaagga cccccaatta agacaggcct tcttgatctt actgaattga tgaaatccca
agatttactt aaccttaatt atgacaatat aaggaatagc aacctgatat ccttttcaaa
attgatttgt tgtgaagggt caaataatat aaatgatggc ttagagttcc tatctgatga
tcctatgaac tttacagaag gcgaagcaat acattcaaca ccaattttta atatatatta
ctcaaagaag ggagataaac atatgacata cagaaatgca atcaaactac tgatagagag
ggaaactaag atttttgaag aagcctttac attcagcgag aatggcttca tatcaccgga
gaatcttggt tgcttagaag ctgtggtatc attaataaaa ttgttgaaaa ctaatgaatg
gtccacagtt atagacaaat gtatccatat atgcttaata aaaaatggca tggaccatat
gtatcattca tttgatgttc ccaaatgctt tatggggaat cctattacta gagatataaa
ttgggtaatg tttagagaat ttatcaatag cttaccagga acagatatac cgccatggaa
tgttatgaca gagaacttta agaagaaatg tattgcccta ataaattcca aatttgaaac
acagagagat ttctcagaat tcactaaatt gatgaaaaag gaaggtggaa ggagcaatat
agagtttgac tagtagttat gggtttatag gggactgaca attagaccat aaatctggga
ggggtttgaa attggctaag attcaaaaag aggggaatta acagcaactg tataaatttg
tagagtggag cacactact3′

Toscana virus partial S segment RNA isolated from an American Tourist returning from Sicily, 354bp (GU799570)

5’tttgcttatc aaggatttga cccaaagcgc attgttcagt tggtcaagga gagaggaaca

gcaaagggca gggattggaa gaaagatgtg aagatgatga ttgtgctgaa ccttgtcagg

gggaacaagc cagaggccat gatgaagaaa atgtcagaga agggtgcttc tattgtggcc

aacctgattt cagtctatca gctgaaagaa gggaatcctg gcagggacac tatcactctg

tcaagggtgt cagctgcatt tgttccgtgg actgttcagg cactacgtgt cctgtcagaa

tccctgcctg tttctgggac caccatggat gcca3’

Toscana virus partial M segment RNA isolated from an American Tourist returning from Sicily, 424bp (GU799571)

5’acaatcatta ctaccctgct catggccagt gcagtaatgg caggaccttt ggagaacagg

gaaacaaacc acttactaaa caggccaggg aatggagctt acaccctaag tgactttgct

gagtccacct gcacgctagc ttatggctca gagtgcaaat cctgggagca ccaactggat

gagctgtctt ttcccttctt ccactccaac cttgacaagt acagcatgct ggaggctgca

acagagacaa taccgatact caacaagagc tcagctgtct gcacaatttc cccatccacg

cactcatcaa atgcatgtgg tagggaagcc agcctcataa agaagaaatg tggaagcaat

atgtcagctt tcttttatgt taacctagcc ggtcagataa ctgttgtcaa gtgtgataca

aacc3’