The Dissertation Committee for Maria Natividad B. Cajimat certifies that this is the approved version of the following dissertation:

GENETIC DIVERSITY AND TAXONOMICAL RELATIONSHIPS AMONG THE TACARIBE SEROCOMPLEX VIRUSES (FAMILY *ARENAVIRIDAE*)

Committee:

Charles F. Fulhorst, D.V.M., Dr.P.H., Supervisor

Joan E. Nichols, Ph.D.

Anthony Simmons, M.D., Ph.D.

Thomas G. Wood, Ph.D.

Stuart T. Nichol, Ph.D.

Dean, Graduate School

GENETIC DIVERSITY AND TAXONOMICAL RELATIONSHIPS AMONG THE TACARIBE SEROCOMPLEX VIRUSES (FAMILY *ARENAVIRIDAE*)

by

Maria Natividad B. Cajimat, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas Medical Branch

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas Medical Branch

September 2007

Key words: *Arenaviridae*, Tacaribe serocomplex, genetic diversity, taxonomy

Dedication

To my family in the Philippines and in Texas.

Acknowledgements

I am eternally grateful to Dr. Charles F. Fulhorst for being an excellent educator and scientist and a patient mentor. I have learned a great deal from Dr. CF throughout my career in graduate school – from classical virological techniques, biological safety in working with viruses, writing grants and manuscripts, to as simple as getting excited about science. I came to his lab knowing very little about arenaviruses and hantaviruses. Dr. CF gave me the opportunity to find out for myself what they were and what they can do. I often struggled while working with these viruses. His support for my work was often in the form of the following quote: "If it was easy, everybody would do it". I will never have enough words to express my gratitude to him for his continuous support of my career as a scientist. There were students who thought Dr. CF was one of the most critical and overly strict mentors on campus, and these are my words to them – "You gave up on a one in a million chance to work with an intelligent and talented scientist". "Thank you" Dr. CF for what you've done for me as a person, my career, and my family.

I am forever indebted to Ms. Mary Louise Milazzo for being a great teacher, a constant listener, and a good friend inside and outside the lab. I have learned the best cell culture techniques and safest way of working with viruses and laboratory animals from Mary Lou. I thank Mary Lou for the constant encouragement and support she has given me in the lab. The work that I have done would not be possible without her efforts – "thank you" for working endless hours in the BSL-4 lab at the Centers for Disease Control and Prevention to propagate the strains of South American hemorrhagic fever arenaviruses. "Thank you" Mary Lou for all that you've done for my career but more importantly, for your love and support throughout the years.

I want to acknowledge Dr. CF's colleagues - Dr. Pierre E. Rollin, Dr. Thomas G. Ksiazek, and Dr. Stuart T. Nichol of the Special Pathogens Branch at the Centers for Disease Control and Prevention for their continuous support of my scientific work on arenaviruses and hantaviruses. "Thank you" for endorsing the project and enabling Mary Lou to work in the BSL-4 lab and access the arenavirus collection and other laboratory resources. This work would not be possible without your support.

I would like to extend my deepest gratitude to my dissertation committee - Dr. Joan Nichols, Dr. Tony Simmons, Dr. Tom Wood, and Dr. Stuart Nichol – for their encouragement and patience throughout my graduate education. I thank them for their critical comments and helpful suggestions on this dissertation and their continuous support of my career. I also want to acknowledge the late Dr. Robert E. Shope who was a member of my earlier dissertation committee for his interest in my research.

I am thankful to Ms. Martha Lewis and Dr. Kley Hughes in the Microbiology and Immunology graduate program for their encouragement throughout the years I stayed in the program. I want to acknowledge Dr. Robert Bradley of Texas Tech University for his helpful suggestions in conducting phylogenetic analyses, valuable insight into the phylogeography of the North American cricetid rodents, and completing some of the maximum likelihood analyses on his computer. I also want to acknowledge Mr. Steve

Smith and Ms. Esther Surriga of the Protein Chemistry Laboratory. The sequencing of the arenaviruses would not have been completed without their help. I thank them for allowing me to run my experimental sequencing reactions.

I want to acknowledge my closest friend in the program and graduate school, Dr. Nanchaya Wanasen. I will never find another friend and roommate as caring, and as patient as her. Being away from my family was difficult at the beginning and Nanchaya was the only "family" I had during the first year in graduate school. I also want to express my heartfelt gratitude to the first friends I met in Pathology who were post-doctoral scientists at the time – Drs. Sophie LePogam, Pei-Ching Tai, John Paul Mutebi, and Rene Rynbrand. I thank them for their continued support of my graduate training and scientific work.

I want to acknowledge my colleagues and friends in the field – Dr. Mario Santiago, Dr. Julius Hafalla, Mr. Keng Pineda, Dr. Gen Hernandez, Ms. Judith Alamares, Ms. Edel Segubre, Dr. Homer Pantua – who have always encouraged and inspired me to pursue graduate studies. I want to acknowledge my Physics teacher and friend, Mr. Rex Absin, not only for his encouragement and support of my career but also for being an excellent example of a good graduate student and parent.

This dissertation and other work related to the research described herein was supported by funding from the National Institutes of Health to Dr. Charles F. Fulhorst - NIH grants AI-41435 entitled "Ecology of emerging arenaviruses in the southwestern United States" and AI-53428 entitled "Rapid, accurate diagnostic assays for arenaviral infections".

I would not have completed my graduate training if not for the love and support of family in the Philippines and here in Galveston. I thank my loving parents for their endless support and sacrifices, for allowing me to pursue this endeavor eventhough it meant not seeing one another and not being there for each other in good times and in bad times. I want to express my heartfelt appreciation to my husband, Mr. Brigge Dondonay, and my extended family on the island, for the love, support, encouragement, and prayers. I thank my little one, Joseph Gabriel, for being a happy baby. He's a source of joy and encouragement in the midst of many challenges.

GENETIC DIVERSITY AND TAXONOMICAL RELATIONSHIPS AMONG THE TACARIBE SEROCOMPLEX VIRUSES (FAMILY *ARENAVIRIDAE*)

Publication No._____

Maria Natividad B. Cajimat, Ph.D.

The University of Texas Medical Branch, 2007

Supervisor: Charles F. Fulhorst

The three objectives of this dissertation were to 1) examine the natural host relationships and origins of the North American Tacaribe serocomplex viruses (family *Arenaviridae*), 2) extend our knowledge of the genetic diversity of the Tacaribe serocomplex viruses, and 3) establish genetic criteria for species demarcation within the *Arenaviridae*. Study 1 -- The results indicated that *Neotoma macrotis* (large-eared woodrat) is the principal host of *Bear Canyon virus*, the present-day association of *Bear Canyon virus* with *Peromyscus californicus* (California mouse) represents a successful host-jumping event from *N. macrotis* to *P. californicus*, and the genomes of the North American arenaviruses are not a product of homologous genetic recombination between ancestors of the present-day South American arenaviruses. Study 2 --The nucleotide sequences of the small genomic segments of 12 strains of *Junin virus*, 7 strains of *Machupo virus*, and 11 strains of *Guanarito virus* were determined to improve our knowledge of strain-to-strain genetic variation within these South American arenavirus species. This new knowledge will enable development of rapid, accurate (sensitive and specific) sequence-based assays for detection of arenavirus-specific RNA in acute-phase clinical specimens from persons affected by South American arenaviral hemorrhagic fevers. Study 3 -- The Eighth Report of the International Committee on Taxonomy of Viruses set forth 4 criteria for demarcation of a species within the *Arenaviridae*: i) the virus must occupy a unique ecological niche, ii) the virus must represent a unique serotype, iii) whether the virus is pathogenic for humans, and iv) the virus must exhibit significant amino acid sequence differences from strains of other arenaviral species. The purpose was to define "significant amino acid difference" in the context of species demarcation. Pairwise comparisons of the amino acid sequences of the glycoprotein precursors and nucleocapsid proteins of 70 arenaviruses, representing the 22 species in

the *Arenaviridae*, indicated that the minimum nonidentity between strains of different species was 15.8% and 11.2%, respectively. Thus, arenaviruses that exhibit at least 15.8% difference in the glycoprotein precursor and at least 11.2% difference in the nucleocapsid protein may be strains of a novel arenavirus species.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

FAMILY *ARENAVIRIDAE*

History and taxonomy

The virus family *Arenaviridae* comprises the genus *Arenavirus*, two serocomplexes, and 22 species (Salvato *et al*., 2005). The lymphocytic choriomeningitis-Lassa (Old World) complex includes *lymphocytic choriomeningitis virus* (LCMV), *Lassa virus* (LASV), *Ippy virus* (IPPYV), *Mobala virus* (MOBV), and *Mopeia virus* (MOPV). The Tacaribe (New World) complex includes *Bear Canyon virus* (BCNV), *Tamiami virus* (TAMV), and *Whitewater Arroyo virus* (WWAV) in North America, *Tacaribe virus* (TCRV) on Trinidad, and *Allpahuayo virus* (ALLV), *Amapari virus* (AMAV), *Cupixi virus* (CPXV), *Flexal virus* (FLEV), *Guanarito virus* (GTOV), *Junin virus* (JUNV), *Latino virus* (LATV), *Machupo virus* (MACV), *Oliveros virus* (OLVV), *Parana virus* (PARV), *Pichindé virus* (PICV), *Pirital virus* (PIRV), and *Sabiá virus* (SABV) in South America (Table 1.1).

Lymphocytic choriomeningitis virus, the prototypical member of the *Arenaviridae*, was discovered in 1933 in the United States. The Armstrong strain of this species was isolated from autopsy tissue of a fatal case that was presumed to be St. Louis encephalitis (Armstrong and Lillie, 1934). Strains of LCMV subsequently were isolated from laboratory white mice (*Mus musculus*), wild mice (*M. musculus*), and Syrian hamsters (*Mesocricetus auratus*) (Gregg, 1975; Rivers and Scott, 1934; Traub, 1935;). *Lymphocytic choriomeningitis virus* is a cause of aseptic meningitis and congenital anomalies in humans (Barton *et al*., 1995; Jahrling and Peters, 1992). Outbreaks of LCMV infections in the past were linked to infected mice and Syrian hamsters that were use for research or sold as pets, cells lines derived from tumors harvested from infected

rodents, or solid organ tissues obtained from an infected donor (Gregg, 1975; CDC, 2005).

Lassa virus was discovered in 1969 in association with an outbreak of hemorrhagic illness among missionary workers in northern Nigeria (Buckley and Casals, 1970; Frame *et al*., 1970). Lassa fever is an important cause of morbidity and mortality in the African countries of Guinea, Liberia, Nigeria, and Sierra Leone (McCormick *et al*., 1987). *Mobala virus*, MOPV, and IPPYV were discovered during surveillance for LASV and other viral zoonoses in different areas of Africa (Gonzalez *et al*., 1983; Johnson *et al*., 1981; Swanepoel *et al*., 1985; Wulff *et al*., 1977).

Tacaribe virus, the prototypical member of the Tacaribe serocomplex, was discovered in 1956. Nineteen strains of TCRV were isolated from the salivary glands and other tissues of frugivorous bats (*Artibeus* spp.) captured on the island of Trinidad during surveillance for rabies virus in the period from 1956 to 1958 (Downs *et al*., 1963). *Junin virus*, the causative agent of Argentine hemorrhagic fever (AHF), was isolated in 1958 from a fatal case of hemorrhagic fever that died in the Buenos Aires Province in northern Argentina (Parodi *et al*., 1958). *Machupo virus*, the etiologic agent of Bolivian hemorrhagic fever (BHF) was isolated in 1963 from a fatal case of hemorrhagic fever that died in the Beni Department in northeastern Bolivia (Johnson *et al*., 1965). The New World arenaviral species identified after the 1960's were discovered during studies on the ecology of arthropod-borne viruses and other viral zoonoses, in response to public health concerns of the emergence of viral diseases such as Yellow fever and St. Louis encephalitis.

Table 1.1. The 22 species in the family *Arenaviridae*

Species[a]	Reservoir host (common name)[b]	Geographic distribution	Year isolated[c]
Lymphocytic choriomeningitis-Lassa serocomplex			
LCMV	*Mus musculus* (house mouse)	Americas, Europe	1933
IPPYV	*Arvicanthis* spp. (grass rat)	Central African Republic	1970
LASV	*Mastomys natalensis* (natal multimammate mouse)	West Africa	1969
MOBV	*Praomys* spp.	Central African Republic	1983
MOPV	*Mastomys natalensis* (natal multimammate mouse)	Mozambique, Zimbabwe	1977
Tacaribe serocomplex			
BCNV	*Peromyscus californicus* (California mouse)	USA (California)	1998
TAMV	*Sigmodon hispidus* (hispid cotton rat)	USA (Florida)	1965
WWAV	*Neotoma albigula* (white-throated woodrat)	USA (New Mexico)	1993
ALLV	*Oecomys bicolor* (bicolored arboreal rice rat)	Peru	1997
AMAV	*Neacomys guianae* (Guiana bristly mouse)	Brazil	1964
CPXV	*Oryzomys capito* (large-headed oryzomys)	Brazil	1964
FLEV	*Oryzomys* spp.	Brazil	1975
GTOV	*Zygodontomys brevicauda* (short-tailed cane mouse)	Venezuela	1990
JUNV	*Calomys musculinus* (drylands vesper mouse)	Argentina	1958
LATV	*Calomys callosus* (large vesper mouse)	Bolivia	1965
MACV	*Calomys callosus* (large vesper mouse)	Bolivia	1963
PARV	*Oryzomys buccinatus* (Paraguayan oryzomys)	Paraguay	1965
PICV	*Oryzomys albigularis* (Tomes' oryzomys)	Colombia	1965
PIRV	*Sigmodon alstoni* (Alston's cotton rat)	Venezuela	1994
OLVV	*Bolomys obscurus* (dark bolo mouse)	Argentina	1988
SABV	Unknown	Brazil	1990
TCRV	*Artibeus* spp.(frugivorous bats)	Trinidad	1956

[a]LCMV, *lymphocytic choriomeningitis virus*; LASV, *Lassa virus*; IPPYV, *Ippy virus*; MOBV, *Mobala virus*; MOPV, *Mopeia virus*; BCNV, *Bear Canyon virus*; TAMV, *Tamiami virus*; WWAV, *Whitewater Arroyo virus*; ALLV, *Allpahuayo virus*; AMAV, *Amapari virus*; CPXV, *Cupixi virus*; FLEV, *Flexal virus*; GTOV, *Guanarito virus*; JUNV, *Junin virus*; LATV, *Latino virus*; MACV, *Machupo virus*; OLVV, *Oliveros virus*; PARV, *Parana virus*; PICV, *Pichindé virus*; PIRV, *Pirital virus*; SABV, *Sabiá virus*; TCRV, *Tacaribe virus*.
[b]The principal hosts of IPPYV, MOBV, FLEV, and SABV are not known.
[c]The year in which the virus-positive specimen was collected or the virus was isolated.

The serological relatedness among TCRV, JUNV, and LCMV was first noted in the late 1960's. Work done by the late Dr. Jordi Casals (1965) established that TCRV and JUNV were antigenically closely related in complement fixation (CF) tests (Casals, 1965; Mettler *et al*., 1963). Studies using the indirect fluorescent antibody (IFA) technique

showed a one-way cross-reaction between LCMV and AMAV, JUNV, LATV, MACV, PARV, PICV, TAMV, and TCRV (Rowe *et al*., 1970b). Subsequent work revealed that members of the Tacaribe serocomplex that were antigenically closely related in the CF test often did not cross-react in neutralization tests (Calisher *et al*., 1970; Casals *et al*., 1975) and LASV cross-reacts with LCMV in the CF test (Casals *et al*., 1975). The work of Dr. Frederick Murphy and colleagues (1969) demonstrated that MACV, TCRV, and LCMV shared a characteristic morphology in thin-section electron microscopy. Various types of cell infected with MACV, TCRV, or LCMV all similarly produced spherical virus particles characterized by club-shaped surface projections and electron-dense internal granules. Together, the results of immunological studies and structural studies provided the foundation for the establishment of a distinct taxonomical group, the Arenoviruses (Murphy *et al*., 1969; Rowe *et al*., 1970b).

The family *Arenaviridae* (Arenovirus taxon) was proposed as a distinct entity in 1970 and initially comprised 10 viruses: LCMV, LASV, AMAV, LATV, JUNV, MACV, PARV, PICV, TAMV, and TCRV (Murphy, 1975; Rowe *et al*., 1970a). The family name was derived from "*areno*", Latin for sand, and comes from the characteristic sandy or granular appearance of arenavirus particles as examined by electron microscopy (Rowe *et al*., 1970a).

Historically, serologic methods, including the CF test and the plaque-reduction neutralization test, were used to define taxonomical relationships among arenaviruses. However, the extreme biohazardous nature of some arenaviruses (i.e., LASV, JUNV, MACV, GTOV, and SABV) has restricted work to the few biosafety level 4 (BSL-4) laboratories in the world. Since 1996, there has been an increasing reliance on the use of nucleotide sequence data to define taxonomical relationships among the arenaviruses (Salvato *et al*., 2005). The results of phylogenetic analyses of partial and complete gene sequence data have been consistent with the results of CF tests and other antibody-antigen binding assays (Bowen *et al*., 1997). The results have also been consistent with separation of the arenaviruses into different species based on the results of *in vitro*

neutralization assays, which usually are species-specific[UTMB1]. The separation of the members of the *Arenaviridae* into 2 major phylogenetic clades (Old World and New World) was consistent with the existence of 2 serocomplexes established by various serologic methods (Bowen *et al*., 1996, 1997).

The *Eighth Report of the International Committee on Taxonomy of Viruses* (ICTV) set forth 5 criteria for species demarcation within the *Arenaviridae*: (1) "association with a distinct host species or group of species", (2) "presence in a defined geographic location", (3) "etiologic agent (or not) of human disease", (4) "significant antigenic differences (including lack of cross-neutralization activity)", and (5) "significant amino acid sequence differences" (Salvato *et al*., 2005). Following the definition of a viral species as "a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche" (van Regenmortel, 2000), a distinct arenavirus must be unique in any 2 or more of these properties.

Natural host relationships of the arenaviral species

Specific members (usually 1 or 2 closely related species) of the rodent family Muridae (Musser and Carleton, 2005) are the natural reservoirs (principal hosts) of the Old World arenaviruses for which natural host relationships have been well characterized. Similarly, specific members of the rodent family Cricetidae (Musser and Carleton, 2005) are the natural reservoirs of the New World arenaviruses for which natural host relationships have been well characterized (Table 1.1). For example, the ubiquitous house mouse (*M. musculus*) is the principal host of LCMV (Childs and Peters, 1993) and the drylands vesper mouse (*Calomys musculinus*) in northern Argentina is the principal host of JUNV (Mills *et al*., 1992). The single exception to the close association between the order Rodentia and the *Arenaviridae* may be TCRV, which has been isolated only from frugivorous bats (*Artibeus* spp.) captured on Trinidad (Downs *et al*., 1963).

The hallmark of the arenaviruses is their ability to establish chronic infections in their respective principal rodent hosts. The chronic carrier state in individual rodents appears to be critical to the long-term maintenance of arenaviruses in nature. The results of previous laboratory studies with MACV in its principal host, the large vesper mouse (*C. callosus*), suggested that chronic infection can negatively affect the fitness of individual rodents or the ability to pass their genes on to the next generation (Childs and Peters, 1993). Conceptually, any negative effect on host fitness could affect the ability of a rodent species to serve as a reservoir.

The results of phylogenetic studies suggested that the present-day host associations of some arenaviruses are a product of a long-standing relationship between the *Arenaviridae* and the order Rodentia (Bowen *et al*., 1996, 1997). Evidence for this ancient virus-rodent host relationship includes the association of phylogenetically closely related arenaviruses with phylogenetically closely related rodent species. For example, JUNV with the drylands vesper mouse (*C. musculinus*) in Argentina and MACV with the large vesper mouse (*C. callosus*) in Bolivia (Childs and Peters, 1993). Other examples include the associations of FLEV, PARV, and PICV with specific members of the rodent genus *Oryzomys* (Clegg, 2002). These 3 South American arenaviruses were monophyletic in sequence analyses (Bowen *et al*., 1997).

The geographic distribution of an arenaviral species subsumes the geographic range of its principal host. For example, TAMV is known only to occur in southern Florida (Calisher *et al*., 1970; Jennings *et al*., 1970). However, the geographic range of its principal host, the hispid cotton rat (*S. hispidus*), extends from Iowa and Nebraska southward through Mexico and into Central America (Hall and Kelson, 1959). The occurrence of human disease caused by an arenavirus usually is limited to small area(s) within the geographic range of the virus and its principal host (Childs and Peters, 1993; Mills *et al*., 1992; Salazar-Bravo *et al*., 2002). For examples, strains of GTOV have been isolated from short-tailed cane mice captured in 5 different states (Apure, Barinas, Cojedes, Guarico, and Portuguesa) in western Venezuela (Weaver *et al*., 2000); however,

Venezuelan hemorrhagic fever appears to be restricted to a small area that includes southern Portuguesa and rural areas in Barinas near the Barinas-Portuguesa border (de Manzione *et al*., 1998).

It is assumed that humans usually become infected with arenaviruses via direct contact with infected rodents (mucosal or cutaneous exposure) or inhalation of infectious virus in aerosolized droplets of urine, saliva, respiratory secretions, or feces from infected rodents. The severity of disease caused by an arenavirus can range from mild to fatal. The occurrence of human-to-human transmission of MACV has been documented in a household setting (CDC, 1994) and human-to-human transmission of LASV has been documented in both hospital and household settings (McCormick and Fisher-Hoch, 2002). The risk of arenaviral infection in humans depends upon a combination of factors which include human behavior (occupational or recreational activities), ecological factors that influence infectivity of rodents, and other variables that affect the frequency and intensity of human exposure to infected rodents (Mills and Childs, 1998).

Human health significance of the arenaviruses

Six of the 22 arenaviral species have been causally associated with human disease. As noted previously, LCMV is an agent of acute central nervous system disease and congenital anomalies (Jahrling and Peters, 1992), LASV is the agent of Lassa fever in western Africa. And JUNV, MACV, GTOV, and SABV are agents of hemorrhagic fever in Argentina, Bolivia, Venezuela, and Brazil, respectively (Peters, 2002). Infection with the LASV and the 4 other hemorrhagic fever viruses is characterized by high (10 – 30%) case-fatality rates.

Lassa virus and the 4 South American hemorrhagic fever arenaviruses have been classified by the National Institutes of Health and the Centers for Disease Control and Prevention as Category A priority pathogens. The Category A pathogens are high-priority biological agents that could potentially be used in terrorism or biological warfare, and as

such, capable of disrupting the public health system. The other arenaviruses are not known to be naturally associated with severe human disease. Many of these viruses have only been discovered within the last 10 to 15 years, and the natural history and clinical importance of these viruses have not been thoroughly studied.

Virion structure and genomic organization

Virions of arenaviruses are spherical to pleomorphic, 50- to 300-nm in diameter, and possess a lipid-bilayer envelope. The lipid envelope, derived from infected cells, is covered with club-shaped, spike-like projections (Buchmeier *et al*., 2001). These surface projections are composed of the viral glycoproteins, GP1 and GP2. Arenavirions appear granular due to the presence of variable numbers of host cell-derived ribosomes enclosed within the virion (Murphy *et al*., 1969). Within the virion are ribonucleoprotein complexes or ribonucleocapsids that consist of the single-stranded genomic RNA segments in complex with molecules of the nucleocapsid (N) protein and RNA-dependent RNA polymerase (RdRp). Ribonucleocapsids of arenaviruses are helical, organized into circular structures, and range in length from 400- to 1300-nm (Salvato *et al*., 2005).

The arenavirus genome consists of 2 single-stranded RNA molecules, designated L (large) and S (small) (Southern, 1996). The L segment (~7.2 kb) consists of a 5' noncoding region (NCR), the Z protein gene, an intergenic noncoding region (IR), the RdRp gene, and a 3'NCR. Similarly, the S segment (~3.5 kb) consists of a 5'NCR, the glycoprotein precursor (GPC) gene, an IR, the N protein gene, and a 3'NCR.

Replication strategy

The arenaviruses utilize an ambisense coding strategy (Bishop, 1986). The 2 genes on each segment are separated by an IR and encoded in opposite polarities (Figure 1.1). The RdRp (sometimes referred to as L protein) and N protein, encoded on the 3'

region of the L and S segments, respectively, are translated from genome-complementary sense subgenomic mRNAs transcribed from the genomic RNA segments. The Z protein and GPC, encoded on the 5' region of the L and S segments, respectively, are translated from genome-sense subgenomic mRNAs transcribed from the full-length genome-complementary replicative intermediate (RI). The subgenomic mRNAs for the Z protein and GPC are made only after the initiation of viral RNA replication, that is – synthesis of the full-length RI, which serves as a template for both transcription of the Z protein and GPC mRNAs and replication of the full-length genomic RNAs.

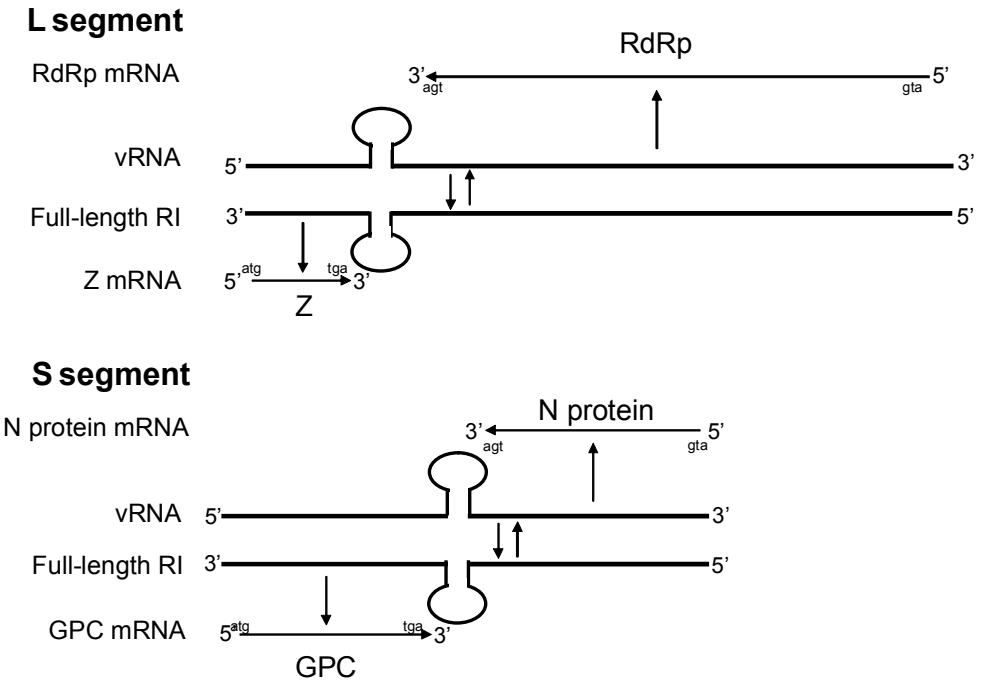


Figure 1.1. Genomic organization and replication strategy of arenaviruses. [Image modified from Meyer *et al*., 2002 Curr Top Microbiol Immunol] L, large genomic segment; RdRp, RNA-dependent RNA polymerase; vRNA, viral genomic RNA; RI, full-length replicative intermediate; Z, zinc-binding protein; S, small genomic segment; N, nucleocapsid; GPC, glycoprotein precursor.

The IR on the L RNA and on the S RNA is rich in G (guanosine) and C (cytosine) nucleotides and is partially complementary along the length of the IR. The IR has the potential to form relatively stable stem-loop or hairpin structures due to this complementarity. The formation of secondary structures in the IR may serve as signal for the termination of transcription (Meyer *et al*., 2002).

The NCR at the 5'- and 3'-ends of the L and S segments are presumed to contain all the *cis*-acting elements needed for initiation of transcription and replication (Meyer *et al*., 2002). The 3'-terminal nucleotide sequences of the L and S RNA segments are very similar and presumably are highly conserved among the arenaviruses (Auperin *et al*., 1982a,b). The sequence of the last 19 nucleotides on the 3'-end (terminus) of the L and S RNA segments is inverse complementary to the sequence of the first 19 nucleotides on the 5'-end of the respective RNA segments (Auperin *et al*., 1982a; Djavani *et al*., 1997). It has been proposed that the sequence complementarity enables the formation of a relatively stable intramolecular duplex (i.e., panhandle structure) that may be critical for the initiation of transcription and replication (Meyer *et al*., 2002).

Virion proteins

As indicated previously, the Z protein (~11 kDa) and the RdRp (~250 kDa) are encoded on the L segment (Buchmeier, 2002). The Z protein is a minor structural component of the virion. It has a RING domain or RING-finger motif that was shown to bind zinc; hence, the name "Z protein". RING domains are known to mediate protein-protein interactions in cells (Borden, 2000). Recent studies demonstrated that the integrity of the RING domain in the Z proteins of LCMV and TCRV was needed for the interaction of the Z protein with the RdRp (Cornu and de la Torre, 2001, 2002; Jacamo *et al*., 2003). The Z protein at the C terminus contains a late domain, a sequence motif that is found in matrix proteins of many RNA viruses. The results of recent studies suggest that the Z protein functions similarly to a matrix protein, for the recruitment and assembly of viral components that gets packaged into virus particles (Perez *et al*., 2003;

Strecker *et al*., 2003). Characteristic of negative sense RNA viruses, the RdRp is a component of ribonucleocapsids and virions (Buchmeier, 2002). After entry and uncoating of the virus in the cytoplasm, the virion-associated RdRp transcribes the genes located on the 3'-end of the RNA segments, the RdRp gene and the N protein gene. The RdRp contains all the enzymatic functions of the virus and includes polymerase, helicase, and methyl transferase activities.

The proteins encoded on the S segment are the GPC (~70 - 80 kDa) and the N protein (~60 - 68 kDa) (Buchmeier, 2002). The GPC is expressed as a polyprotein that is post-translationally cleaved by an SKI-1/S1P protease to yield the envelope glycoproteins, GP1 (40-46 kDa) and GP2 (35 kDa) (Lenz *et al*., 2001; Beyer *et al*., 2003). The first 58 amino acids of the GPC make up the signal peptide which is also cleaved by a host protease (Eichler *et al*., 2003). The GP1 and GP2 glycoproteins form oligomeric complexes on the surface of the lipid envelope. The GP1 is a peripheral membrane glycoprotein exposed on the surface of the virion and is the virion attachment protein that binds to receptors present on host cells (Buchmeier *et al*., 2001). GP2 is a class I fusion protein embedded within the lipid envelope and is responsible for the pH-dependent fusion of the viral lipid envelope with the endosomal membrane within host cells that results in uncoating of the virus in the cytoplasm (Buchmeier, 2002). The N protein is the most abundant structural protein within virions (Buchmeier, 2002). Molecules of the N protein encapsidate the genomic RNA segments and together with the RdRp comprises the ribonucleoprotein complex.

Mechanisms of genetic change

Genetic diversity among the arenaviruses is predominantly generated by genetic drift – the accumulation of point mutations and insertion/deletion mutations (Bowen *et al*., 1997). Point mutations are generated by the lack of a proofreading mechanism in the RdRp of RNA viruses. Sequence differences among strains of a viral species and among

viral species belonging to a phylogenetic lineage are primarily attributed to the accumulation of these mutations (Charrel *et al*., 2002).

There is circumstantial evidence that arenaviruses also evolve through genetic shift – that is, by intra-segmental genetic recombination or genome segment reassortment. The S segment of the North American arenaviruses appears to be a product of recombination between 2 phylogenetically distinct viruses (Archer and Rico-Hesse, 2002; Charrel *et al*., 2002). Specifically, the GPC and the N protein genes of BCNV, TAMV, and WWAV, appear to have been acquired from a virus that was most closely related to AMAV, CPXV, GTOV, JUNV, MACV, SABV, or TCRV and from a virus that was most closely related to ALLV, FLEV, PARV, PICV, or PIRV, respectively. Co-infection of a single host by 2 (or more) distinct viruses is required for genetic recombination to occur.

Genome segment reassortment occurs among viruses with multipartite genomes (i.e., 2 or more genomic segments). The acquisition or exchange of genomic segments occurs also during co-infection of a host by 2 or more distinct viruses. Although it has not been observed in nature, reassortants between strains of the same species and between strains of different species have been achieved in the laboratory setting, as in the case of the WE and Armstrong strains of LCMV and between LASV and MOPV, respectively (Riviere and Oldstone, 1986; Lukashevich, 1992). As evident in other segmented viruses, genome segment reassortment can contribute to the evolution of the arenaviruses.

Phylogenetic history of the arenaviruses

The most comprehensive assessment of the phylogenetic history of the *Arenaviridae* was reconstructed using complete GPC sequences and complete N protein sequences (Archer and Rico-Hesse, 2002; Charrel *et al*., 2002). In both studies, the *Arenaviridae* was divided into 5 phylogenetic lineages – Old World, and New World A, B, C, and North American (also known as recombinant) (see Figure 1.2). In these studies,

the Old World lineage was represented by LCMV, LASV, and MOPV. The New World lineage A was represented by ALLV, FLEV, PARV, PICV, and PIRV, lineage B by AMAV, CPXV, GTOV, JUNV, MACV, SABV, and TCRV, lineage C by LATV and OLVV, and North American by BCNV, TAMV, and WWAV. As mentioned previously, the S segments of BCNV, TAMV, and WWAV appear to be a product of genetic recombination between 2 viruses, that is – a virus phylogenetically most closely related to the New World lineage B viruses was the source of the GPC gene and a virus phylogenetically most closely related to the New World lineage A viruses was the source of the IR and the N protein gene (Charrel *et al*., 2002).

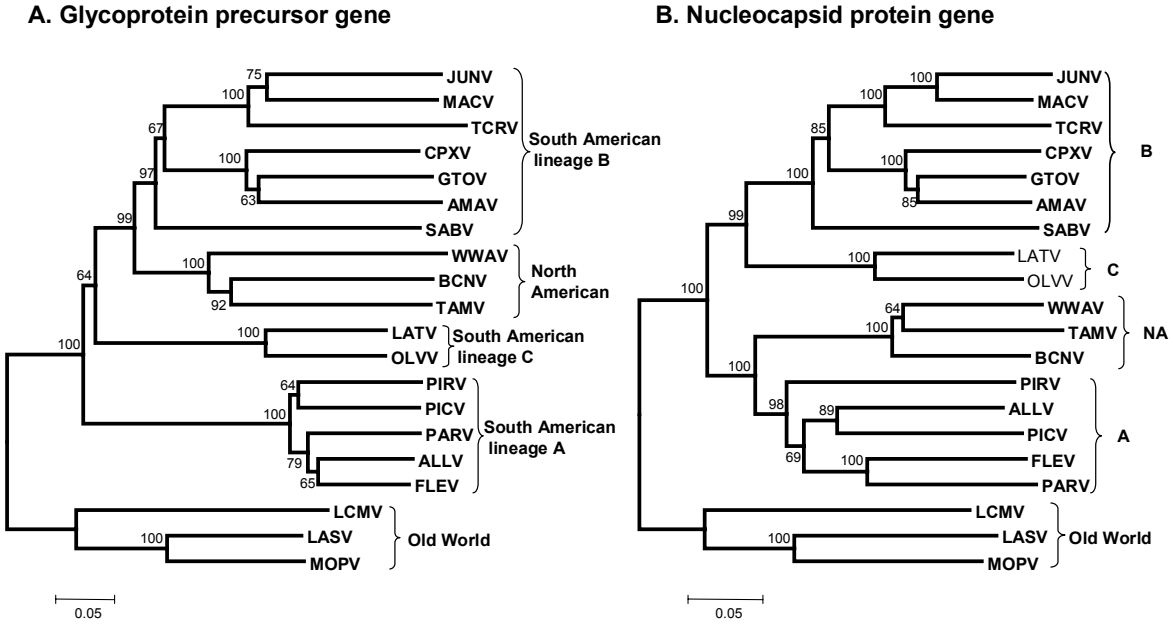


Figure 1.2. Phylogenetic relationships in the *Arenaviridae*. Phylogenetic relationships were reconstructed based on neighbor-joining analyses of complete GPC sequences (A) and complete N protein sequences (B) [modified from Charrel *et al*., 2002 Biochem Biophys Res Comm]

RESEARCH AIMS

The broad objective of the research described herein was to extend our knowledge of the genetic diversity among the arenaviruses native to the New World. This new knowledge could be used to improve our understanding of the phylogenetic history of the New World arenaviruses, develop rapid, accurate nucleotide sequence-based assays for detection of arenavirus-specific RNA in biological specimens and develop gene-specific therapies for the South American arenaviral hemorrhagic fevers. The specific aims were -

Aim 1. Examine further the natural host relationships and the evolutionary history of the North American arenaviral species, that is – BCNV, TAMV, and WWAV
Aim 2. Extend our knowledge of the genetic diversity among the South American hemorrhagic fever arenaviruses
Aim 3. Establish and assess a genetic criteria for species demarcation within the *Arenaviridae*

Chapter 2: Principal host relationships and evolutionary history of the North American arenaviruses

INTRODUCTION

The virus family *Arenaviridae* comprises 2 serocomplexes and 22 species (Salvato *et al*., 2005). The Tacaribe (New World) complex includes *Bear Canyon virus* (BCNV), *Tamiami virus* (TAMV) and *Whitewater Arroyo virus* (WWAV) in North America (see Figure 2.1), *Tacaribe virus* (TCRV) on Trinidad in the Caribbean Sea, and *Allpahuayo virus* (ALLV), *Amapari virus* (AMAV), *Cupixi virus* (CPXV), *Flexal virus* (FLEV), *Guanarito virus* (GTOV), *Junin virus* (JUNV), *Latino virus* (LATV), *Machupo virus* (MACV), *Oliveros virus* (OLVV), *Parana virus* (PARV), *Pichindé virus* (PICV), *Pirital virus* (PIRV) and *Sabiá virus* (SABV) in South America. The lymphocytic choriomeningitis-Lassa (Old World) complex includes *Ippy virus* (IPPYV), *Lassa virus* (LASV), *lymphocytic choriomeningitis virus* (LCMV), *Mobala virus* (MOBV) and *Mopeia virus* (MOPV).

Specific members of the subfamilies Neotominae and Sigmodontinae in the rodent family Cricetidae (Musser and Carleton, 2005) are the principal hosts of the New World arenaviral species for which natural host relationships have been well characterized. For example, the hispid cotton rat (*Sigmodon hispidus*) in southern Florida is the principal host of TAMV (Calisher *et al*., 1970; Jennings *et al*., 1970), the white-throated woodrat (*Neotoma albigula*) in northwestern New Mexico is the principal host of WWAV (Fulhorst *et al*., 1996) and Alston's cotton rat (*Sigmodon alstoni*) in western Venezuela is the principal host of PIRV (Fulhorst *et al*., 1997, 1999).

The California mouse (*Peromyscus californicus*) in southern California is a natural host of BCNV (Fulhorst *et al*., 2002a). The BCNV prototype strain AV A0070039 was isolated from a California mouse captured in November 1998 near Bear Canyon in the Santa Ana Mountains in southern California (Fulhorst *et al*., 2002). Bear

Canyon is located in western Riverside County near the Riverside County-Orange County line. Strains of BCNV subsequently were isolated from 4 (19.1%) of 21 California mice captured in June 1998 at 2 sites in the Santa Ana Mountains in eastern Orange County (Fulhorst *et al*., 2002a).

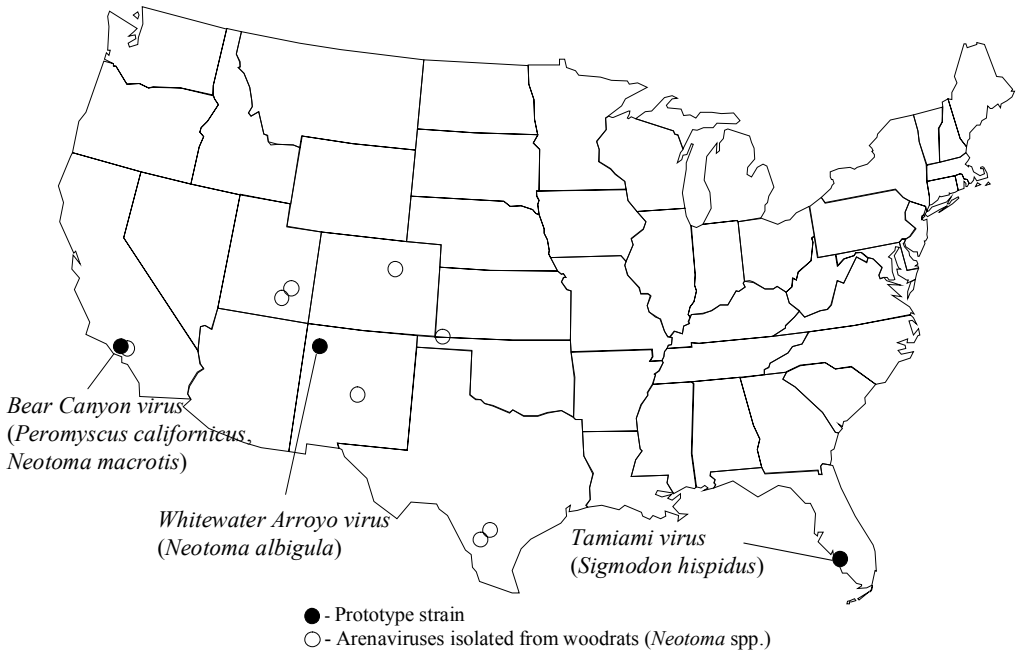


Figure 2.1 Geographic distribution and natural host relationships of the North American arenaviruses. *Bear Canyon virus* strains AV A0060209 and AV A0070039 were isolated from California mice (*Peromyscus californicus*) captured in 1998 in southern California. Strain AV 98470029 was isolated from a large-eared woodrat (*Neotoma macrotis*) captured in 1998 in southern California. The principal host of *Tamiami virus* in southern Florida is the hispid cotton rat (*Sigmodon hispidus*). The principal host of *Whitewater Arroyo virus* in New Mexico is the white-throated woodrat (*Neotoma albigula*). Arenaviruses phylogenetically closely related to *Whitewater Arroyo virus* were isolated from white-throated woodrats (*N. albigula*) in New Mexico and Oklahoma, Mexican woodrats (*N. mexicana*) in Colorado, New Mexico, and Utah, bushy-tailed woodrat (*N. cinerea*) in Utah, and southern plains woodrat (*Neotoma micropus*) in southern Texas.

In previous studies (Bennett *et al*., 2000; Fulhorst *et al*., 2002a), antibody to an arenavirus was found in 1 (50.0%) of 2 dusky-footed woodrats (*Neotoma fuscipes*) and 4 (50.0%) of 8 California mice captured at a site in western Riverside County, 5 (16.1%) of 31 dusky-footed woodrats and 7 (21.9%) of 32 California mice captured at 7 sites in

eastern Orange County, and 3 (6.8%) of 44 dusky-footed woodrats and 1 (4.2 %) of 24 California mice captured at 14 sites in southern Orange County. Note that *Neotoma macrotis* (large-eared woodrat) in Riverside County, Orange County, Los Angeles County, and several other counties in southern California recently was elevated from subspecific status within the *N. fuscipes* species complex (Matocq, 2002). Thus, the antibody-positive woodrats in the previous studies likely were large-eared woodrats, not dusky-footed woodrats.

Studies done in the 1990's established that multiple (different) arenaviruses coexist in certain regions of South America. For example, JUNV and OLVV are sympatric in the pampas of northern Argentina (Mills *et al*., 1996) and GTOV and PIRV are sympatric on the plains of western Venezuela (Fulhorst *et al*., 1997, 1999). The first objective of the present study was to determine the identity of the arenavirus associated with the large-eared woodrat in the Santa Ana Mountains. Hypothetically, the arenavirus associated with the large-eared woodrat in the Santa Ana Mountains is different from BCNV.

Independent analyses of complete GPC sequences and complete N protein sequences in a previous study (Charrel *et al*., 2002) delineated 5 phylogenetic lineages within the *Arenaviridae*: North American (BCNV, TAMV, and WWAV), South American lineage A (ALLV, FLEV, PARV, PICV, and PIRV), South American lineage B (AMAV, CPXV, GTOV, JUNV, MACV, SABV, and TCRV), South American lineage C (LATV and OLVV), and Old World (LASV, LCMV, and MOPV). Note that the analysis of the GPC sequence data placed the North American lineage in a sister relationship to the South American lineage B whereas the analysis of the N protein sequence data placed the North American lineage in a sister relationship to the South American lineage A. Also note that the sister relationship between the North American lineage and the South American lineage B in the analysis of the GPC sequence data and the sister relationship between the North American lineage and the South American

lineage A in the analysis of the N protein sequence data were strongly supported by bootstrap analyses.

Previously, our most comprehensive knowledge of the evolutionary history of the arenaviral L genomic segment was based on analyses of amino acid sequences predicted from a small fragment (288- to 300-nt) of the RdRp genes of BCNV strain AV A0060209, TAMV strain W-10777, WWAV strain AV 9310135, TCRV strain TRVL II573, 11 South American arenaviruses and 3 Old World arenaviruses (Charrel *et al*., 2003). In a neighbor-joining analysis of the RdRp sequence data, the 3 North American arenaviruses were monophyletic, the North American lineage was sister to a lineage that comprised ALLV, PICV and PIRV, and the BCNV-TAMV-WWAV-ALLV-PICV-PIRV lineage was sister to a lineage that comprised AMAV, CPXV, GTOV, JUNV, MACV, SABV and TCRV. Thus, the L segments (RdRp genes) and the 3' halves of the S segments (N protein genes) of the North American arenaviruses appear to be descended from an ancestor of the South American lineage A that emerged after the divergence of the South American lineage A from the South American lineage B. The second objective of this study was to extend our knowledge of the phylogenetic history of the North American arenaviruses, specifically -- to determine whether the L genomic segments of BCNV, TAMV and WWAV are a product of homologous recombination between arenaviruses with significantly different phylogenetic histories.

MATERIALS AND METHODS

The nucleotide sequence of a 3290-nt fragment of the S segment of BCNV strain AV 98470029, a 3301-nt fragment of the S segment and a 7064-nt fragment of the L segment of the BCNV prototype strain AV A0070039 (Fulhorst *et al*., 2002a), a 7103-nt fragment of the L segment of the TAMV prototype strain W-10777 (Calisher *et al*., 1970), a 7094-nt fragment of the L segment of the WWAV prototype strain AV 9310135 (Fulhorst *et al*., 1996), and a 7008-nt fragment of the L segment of AMAV strain BeAn

70563 (Pinheiro *et al*., 1966) were determined in this study. The S segment sequences each extend from within the 5' NCR through the 3'end of the N protein gene. Similarly, the L segment sequences each extend from within the 5' NCR through the 3'end of the RdRp gene. The sequence of the L segment of AMAV strain BeAn 70563 was determined to improve the representation of the South American arenaviral species in this study.

Virus assay

The spleen of a large-eared woodrat (*N. macrotis*) and the spleens of 6 cactus mice (*Peromyscus eremicus*) were tested for infectious arenavirus by cultivation in monolayers of Vero E6 cells as described previously (Fulhorst *et al*., 1996). The woodrat and the cactus mice were captured in September 1998 at a site in Riverside County located approximately 17.7 km southeast of Bear Canyon. Strain AV 98470029 was isolated from the spleen of the woodrat. The virus isolation attempts on the spleens of the 6 cactus mice were negative. All work with infectious virus was done in a biosafety level 3 (BSL-3) laboratory located in the Keiller Bldg.

Preparation of RNA and synthesis of first-strand cDNA

Total RNA was isolated from monolayers of Vero E6 cells infected with strain AV 98470029, BCNV strain AV A0070039, TAMV strain W-10777, WWAV strain AV 9310135 or AMAV strain BeAn 70563 using TRIzol® Reagent (Invitrogen Life Technologies, Inc., Carlsbad, CA). Reverse transcription of arenavirus-specific RNA was carried out as described previously (Günther *et al*., 2000; Cajimat and Fulhorst, 2004), using SuperScript II RNase H^- Reverse Transcriptase (Invitrogen Life Technologies, Inc.) in conjunction with 19C-cons (5'- CGCACMGWGGATCCTAGGC-3'). Oligonucleotide 19C-cons is a derivative of oligonucleotide ARE3'-END (Gonzalez *et al*., 1995) and was expected to anneal to the 19-nt fragment at the extreme 3' end of the S segment and to the 19-nt fragment at the extreme 3' end of the L segment (Auperin *et al*., 1982a; Southern,

1996). The nucleotide sequence of the 19-nt fragment at the extreme 5' end of the S segment and the nucleotide sequence of the 19-nt fragment at the extreme 5' end of the L segment presumably are inverse complementary to the nucleotide sequence of the 19-nt fragment at the extreme 3' end of the S segment and the nucleotide sequence of the 19-nt fragment at the extreme 3' end of the L segment, respectively (Auperin *et al*., 1982a; Djavani *et al*., 1997; Southern, 1996). Thus, 19C-cons was expected to prime synthesis of first-strand cDNA from 4 different templates: the 3' end of the S segment, the 3' end of the replicative intermediate synthesized from the S segment, the 3' end of the L segment and the 3' end of the replicative intermediate synthesized from the L segment.

PCR amplification of arenavirus-specific first-strand cDNA and sequencing reactions

The nucleotide sequence of the 3290-nt fragment of the S segment of strain AV 98470029 and the 3301-nt fragment of the S segment of strain AV A0070039 each were determined from 2 overlapping PCR products: S1 and S2. The S1 products and S2 products were synthesized by using Hot Master Taq DNA polymerase (Eppendorf, Westbury, NY) in conjunction with 19C-cons *and* AVNP4 (Fulhorst *et al*., 2002a) and 19C-cons *and* AVNP13 (5'-GTTGTKTCWGGYTCYCTGAA-3'), respectively. Both strands of each PCR product were sequenced directly.

The nucleotide sequence of the 7064-nt fragment of the L segment of BCNV strain AV A0070039 was determined from 4 PCR products. BCNVL1 (5,045-bp) and BCNVL2 (2,196-bp) were generated from the AV A0070039 first-strand cDNA by using the Triple Master PCR System (Eppendorf) in conjunction with 19C-cons *and* AVPOL16 (5'-TCACTTATTAACAGAAGCCC-3') and 19C-cons *and* AVPOL1 (5'-GAAYTTCTCAAAACATTTGATTTG-3'), respectively. The nucleotide sequence of 1 strand of each PCR product was determined directly. The BCNVL1-BCNVL2 sequence comprised the entire 7064-nt fragment of the L segment of AV A0070039. The sequence of the 7064-nt fragment subsequently was determined from BCNVL3 (3,961-bp) and

BCNVL4 (3,792-bp), which were generated from the AV A0070039 L segment first-strand cDNA by using the Triple Master PCR System (Eppendorf) in conjunction with 19C-cons *and* AVPOL33 (5'-AGGTATGATCACCGAAGTAG-3') and 19C-cons *and* AVPOL83 (5'-CCTGTCCATTAAGCCAAGCC-3'), respectively. BCNVL3 and BCNVL4 were cloned in the pGEM®-T Easy vector (Promega Corp., Madison, WI). The sequence of 1 strand of the cloned BCNVL3 was determined. Multiple attempts to propagate plasmids that contained BCNVL4 were unsuccessful; therefore, 1 strand of BCNVL4 was sequenced directly. The single base discrepancy between the BCNVL1-BCNVL2 sequence and the BCNVL3-BCNVL4 sequence was resolved by comparison with the homologous sequence of a 2,196-bp product generated from the AV A0070039 L segment first-strand cDNA by using Triple Master PCR System (Eppendorf) in conjunction with 19C-cons *and* AVPOL1.

The nucleotide sequence of the 7103-nt fragment of the L segment of TAMV strain W-10777 was determined from 2 overlapping PCR products. TAMVL1 (3148-bp) and TAMVL2 (4457-bp) were generated from the first-strand cDNA by using the Triple Master PCR System (Eppendorf) in conjunction with 19C-cons *and* AVPOL90 (5'-TACACATCRAGTGATGATGAGATC-3') and 19C-cons *and* AVPOL100 (5'-CGGTAACCYCTTGAWCCRTCMACCC-3'), respectively. Both strands of TAMVL1 and TAMVL2 were sequenced directly.

The nucleotide sequence of the 7094-nt fragment of the L segment of WWAV strain AV 9310135 was determined from 2 overlapping PCR products. WWAVL1 (3142-bp) and WWAVL2 (4451-bp) were generated by using the Triple Master PCR System (Eppendorf) in conjunction with 19C-cons and AVPOL90 and 19C-cons and AVPOL100, respectively. Both strands of WWAVL1 and WWAVL2 were sequenced directly. In a preliminary analysis, the sequence of the 318-nt fragment of the RdRp gene of WWAV strain AV 9310135 published previously (Charrel *et al*., 2003) and deposited into the GenBank nucleotide sequence database (accession no. AY216516) was only 78% identical to the sequence of the homologous region of WWAVL2. To provide an

assurance that WWAVL2 was from WWAV strain AV 9310135, a 616-nt fragment of the N protein gene was generated from the AV 9310135 first-strand cDNA and the sequence of the PCR product compared to the sequence of the N protein gene of WWAV strain AV 9310135 published previously (Fulhorst *et al*., 1996). The sequence of the 616-nt fragment was identical to the N protein gene sequence in GenBank accession no. U52180. Subsequently, total RNA was isolated directly from the virus stock that was used to infect the Vero E6 cells (see "Preparation of RNA" above) and a 1976-bp PCR product was generated from the L segment (RdRp gene) first-strand cDNA, using oligonucleotides AVPOL96 (5'- AATAAGAGYGTTGTTGTCCC-3') with AVPOL100 to prime synthesis of the PCR product. The sequence of this PCR product was identical to the sequence of the homologous region of WWAVL2. Together, the sequence of the N protein gene generated from the original first-strand cDNA and the sequence of the 1976-nt fragment of the RdRp gene generated from the virus stock indicate that the sequence of WWAVL2 is from WWAV strain AV 9310135.

The nucleotide sequence of the 7008-nt fragment of the L segment of AMAV strain BeAn 70563 was determined from 2 overlapping PCR products. AMAVL1 (3995-bp) and AMAVL2 (4576-bp) were amplified from the AMAV L segment first-strand cDNA by using the Triple Master PCR System (Eppendorf) in conjunction with 19C-cons *and* AVPOL70 (5'-AGATGTTTCGAAGAACAGG-3') and 19C-cons *and* AVPOL77 (5'-GAGAGGTTGTGAAGTATTCC-3'), respectively. One strand of each PCR product was sequenced directly. Subsequently, AMAVL1 and AMAVL2 were cloned in the pGEM®-T Easy vector (Promega Corp.). One strand of the cloned AMAVL1 and AMAVL2 each were sequenced using plasmid- and virus-specific oligonucleotides to prime the sequencing reactions.

The nucleotide sequence of the 3290-nt fragment of the S segment of BCNV strain AV 98470029 and the 3301-nt fragment of the S segment of BCNV strain AV A0070039 were deposited into the GenBank nucleotide sequence database under accession nos. AY924392 and AY924391, respectively. The nucleotide sequence of the

7064-nt fragment of the L segment of strain AV A0070039, the 7103-nt fragment of the L segment of TAMV strain W-10777, the 7094-nt fragment of the L segment of WWAV strain AV 9310135 and the 7008-nt fragment of the L segment of AMAV strain BeAn 70563 were deposited into the GenBank nucleotide sequence database under accession nos. AY924390, AY924393, AY924395 and AY924389, respectively.

Data analysis

The nucleotide sequence of the GPC gene and the N protein gene of strain AV 98470029 were compared to the homologous sequences of BCNV strain A0060209, BCNV strain A0070039, TAMV strain W-10777, WWAV strain AV 9310135 and 8 South American arenaviruses. The nucleotide sequences of the Z protein genes and RdRp genes of BCNV strain A0070039, TAMV strain W-10777 and WWAV strain AV 9310135 were compared to the homologous sequences of the 8 South American arenaviruses. The GPC, N protein, Z protein and RdRp amino acid sequences were aligned independently, using the computer program CLUSTAL W1.7 (Thompson *et al*., 1994). The nucleotide sequences were aligned manually based on the computer-generated amino acid sequence alignments. The GPC, N protein, Z and RdRp gene sequence alignments were 1683, 1734, 333, and 6942 characters in length, respectively. The analyses of the nucleotide sequence alignments were done by using programs in the computer software package PAUP*, version 4.0b10 (Swofford, 2003) and MEGA version 2.1 (Kumar *et al*., 2001). Nucleotide sequence identities were calculated by subtracting uncorrected *p* model distances from 1.0. Bootstrap support (Felsenstein, 1985) for the results of the neighbor-joining (NJ) analyses of *p* model distances and maximum parsimony (MP) analyses of nucleotide sequence data was based on 1000 pseudoreplicate datasets and 500 pseudoreplicate datasets, respectively. The topologies of the NJ phylograms and corresponding MP phylograms were essentially identical (MP phylograms not shown). *Lymphocytic choriomeningitis virus* strain WE (GenBank accession nos. M22138 and AF004519) and LASV strain Josiah (NC_004296 and

NC_004297) were used as outgroup taxa in the NJ analyses to infer the ancestral node among the New World arenaviruses included in this study.

RESULTS

Strain AV 98470029 is an arenavirus that was isolated in this study from a large-eared woodrat (*N. macrotis*) captured in September 1998 in the Santa Ana Mountains in Riverside County. The nucleotide sequences of the GPC and N protein genes of AV 98470029 were compared to the homologous sequences of 4 other North American arenaviruses, TCRV strain TRVL II573, and 7 South American arenaviruses (Table 2.1). Note that the BCNV strains AV A0060209 and AV A0070039 were isolated in a previous study (Fulhorst *et al.*, 2002a) from California mice captured in the Santa Ana Mountains in Orange County and Riverside County, respectively.

In pairwise comparisons, the nucleotide sequence of the GPC gene of AV 98470029 was 97.0%, 98.1%, 64.0%, and 62.5% identical to the nucleotide sequences of the GPC genes of BCNV strain AV A0060209, BCNV strain AV A0070039, TAMV strain W-10777, and WWAV strain AV 9310135, respectively, and less than 57.3% identical to the nucleotide sequences of the GPC genes of TCRV strain TRVL II573 and the 7 South American arenaviruses included in this study (Table 2.2). Similarly, the nucleotide sequence of the N protein gene of AV 98470029 was 96.4%, 98.0%, 71.8%, and 71.9% identical to the nucleotide sequences of the N protein genes of BCNV strain AV A0060209, BCNV strain AV A0070039, TAMV strain W-10777, and WWAV strain AV 9310135, respectively, and less than 63.9% identical to the nucleotide sequences of the N protein genes of TCRV strain TRVL II573 and the 7 South American arenaviruses included in this study (Table 2.2).

Neighbor-joining analyses of the *p* model distances generated from the GPC and N protein gene sequence alignments (all 3 nucleotide positions included in the calculation of the genetic distances) indicated that AV 98470029 is phylogenetically more closely

related to BCNV strains AV A0070039 and AV A0060209 than to TAMV strain W-10777 or WWAV strain AV 9310135 (Figure 2.2). Together, the high level of sequence identity between strain AV 98470029 and strains AV A0060209 and AV A0070039 and the results of the neighbor-joining analyses of the nucleotide sequence data indicate that AV 98470029 is a strain of BCNV.

Table 2.1. Histories of the 15 arenaviruses included in this study

Species[a]	Strain	Isolated from	
		Country (state)	Host[b]
Bear Canyon virus	AV A0060209	USA (California)	*Peromyscus californicus* (California mouse)
Bear Canyon virus	AV A0070039	USA (California)	*P. californicus*
Bear Canyon virus	AV 98470029	USA (California)	*Neotoma microtis* (large-eared woodrat)
Tamiami virus	W-10777	USA (Florida)	*Sigmodon hispidus* (hispid cotton rat)
Whitewater Arroyo virus	AV 9310135	USA (New Mexico)	*Neotoma albigula* (white-throated woodrat)
Pichindé virus	An 3739	Colombia	*Oryzomys albigularis* (Tomes's oryzomys)
Pirital virus	VAV-488	Venezuela	*Sigmodon alstoni* (Alston's cotton rat)
Amapari virus	BeAn 70563	Brazil	*Neacomys guianae* (Guiana bristly mouse)
Guanarito virus	INH-95551	Venezuela	*Homo sapiens* (human)
Junin virus	XJ13	Argentina	*H. sapiens*
Machupo virus	Carvallo	Bolivia	*H. sapiens*
Sabiá virus	SPH 114202	Brazil	*H. sapiens*
Tacaribe virus	TRVL II573	Trinidad	*Artibeus* spp. (frugivorous bats)
Lassa virus	Josiah	Sierra Leone	*Mastomys natalensis* (natal multi-mammate mouse)
Lymphocytic choriomeningitis virus	WE	USA	*Mus musculus* (house mouse)

[a]Strains AV A0060209 and AV A0070039 were isolated from California mice captured in June 1998 in Riverside County and in November 1998 in Orange County, respectively. Strain AV 98470029 was isolated from a large-eared woodrat captured in September 1998 in Riverside County.
[b]The principal hosts of *Guanarito virus*, *Junin virus*, and *Machupo virus* are *Zygodontomys brevicauda* (short-tailed cane mouse), *Calomys musculinus* (drylands vesper mouse), and *Calomys callosus* (large vesper mouse), respectively. The principal host of *Sabiá virus* has not been determined.

Table 2.2. Nucleotide sequence identities in the glycoprotein precursor gene and nucleocapsid protein gene of 13 New World arenaviruses[a]

	Glycoprotein precursor gene (% nucleotide sequence identity)											
Virus(es)[b]	AV 98470029	BCNV	TAMV	WWAV	PICV	PIRV	AMAV	GTOV	JUNV	MACV	SABV	TCRV
AV 98470029	--	97.0--98.1	64.0	62.5	50.6	51.9	55.4	55.6	57.3	57.2	56.7	53.6
BCNV	96.4--98.0	--	63.9--64.3	62.5--62.9	50.4--50.7	51.6--52.0	55.4--55.9	55.6--55.8	56.9--57.1	57.1--57.2	56.6--56.8	52.9--53.2
TAMV	71.8	71.8	--	61.4	49.2	50.0	55.8	55.1	55.6	55.2	56.6	54.2
WWAV	71.9	72.1--72.3	71.7	--	51.1	52.3	53.1	54.2	54.9	57.0	56.0	53.4
PICV	63.4	63.2--63.6	62.1	63.1	--	72.6	51.1	51.2	49.2	48.9	50.2	50.0
PIRV	63.9	63.7--63.9	62.0	62.2	66.7	--	51.5	50.9	50.5	51.0	52.4	52.2
AMAV	57.7	57.3	55.2	56.6	59.4	57.9	--	67.8	58.2	58.7	57.6	56.9
GTOV	57.3	57.5	57.1	57.9	59.2	59.0	75.3	--	57.9	57.9	58.7	56.4
JUNV	56.9	56.6	57.2	56.5	57.6	59.3	68.2	69.6	--	69.1	57.5	66.6
MACV	59.7	59.6--59.8	56.6	58.1	58.6	60.3	67.4	69.4	76.1	--	58.3	62.9
SABV	58.8	58.5--59.0	57.1	58.0	58.5	57.1	66.8	66.0	66.8	67.3	--	57.0
TCRV	57.8	57.9	57.0	57.0	59.0	58.4	67.6	67.7	72.2	71.5	64.8	--
	Nucleocapsid protein gene (% nucleotide sequence identity)											

[a]Sequence identities among the glycoprotein precursor genes and among the nucleocapsid protein genes are listed above and below the diagonal, respectively.

[b]Strain AV 98470029 (GenBank accession no. AY924392); BCNV, *Bear Canyon virus* strains AV A0060209 and AV A0070039 (AF512833 and AY924390, respectively); TAMV, *Tamiami virus* strain W-10777 (AF512828); WWAV, *Whitewater Arroyo virus* strain AV 9310135 (AF228063); PICV, *Pichindé virus* strain An 3739 (NC_006447); PIRV, *Pirital virus* strain VAV-488 (NC_005894); AMAV, *Amapari virus* strain BeAn 70563

(AF512834); GTOV, *Guanarito virus* strain INH-95551 (NC_005077); JUNV, *Junin virus* strain XJ13 (NC_005081); MACV, *Machupo virus* strain Carvallo (NC_005078); SABV, *Sabiá virus* strain SPH 114202 (NC_006317); TCRV, *Tacaribe virus* strain TRVL II573 (NC_004293). The nucleotide sequences of BCNV strains AV A0060209 and AV A0070039 were 97.2% and 96.3% identical in the glycoprotein precursor gene and nucleocapsid protein gene, respectively.

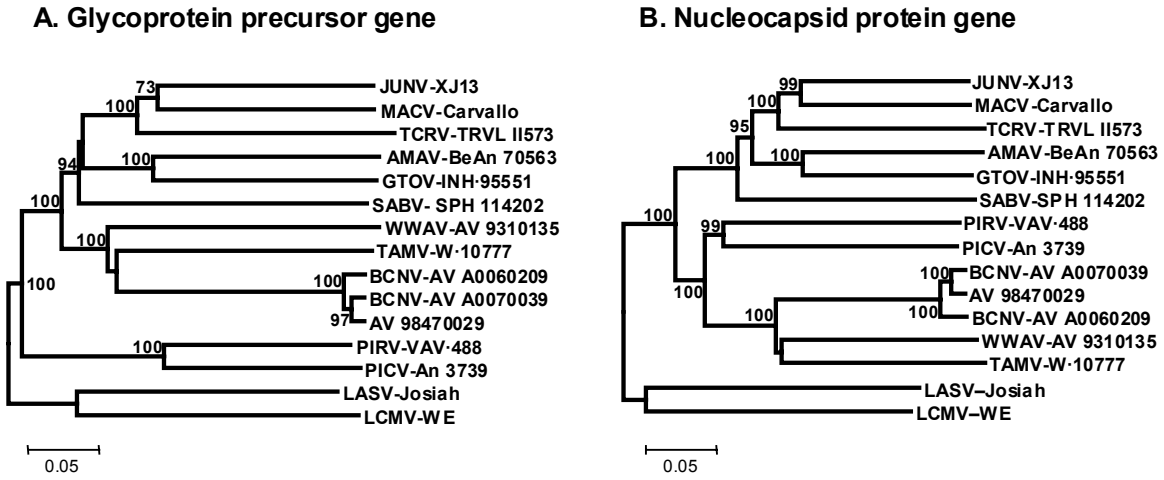


Figure 2.2. Phylogenetic relationships among 15 arenaviruses based on neighbor-joining (NJ) analyses of (A) complete glycoprotein precursor gene sequences and (B) complete nucleocapsid protein gene sequences. Horizontal branch lengths are proportional to nucleotide sequence divergence. The scale bar represents divergence of 0.05. Bootstrap support was based on 1000 pseudoreplicates; values greater than 50% are shown on each interior branch. Strain AV 98470029 (GenBank accession no. AY924392); BCNV, *Bear Canyon virus* strains AV A0060209 and AV A0070039 (AF512833 and AY924390, respectively); TAMV, *Tamiami virus* strain W-10777 (AF512828); WWAV, *Whitewater Arroyo virus* strain AV 9310135 (AF228063); PICV, *Pichindé virus* strain An 3739 (NC_006447); PIRV, *Pirital virus* strain VAV-488 (NC_005894); AMAV, *Amapari virus* strain BeAn 70563 (AF512834); GTOV, *Guanarito virus* strain INH-95551 (NC_005077); JUNV, *Junin virus* strain XJ13 (NC_005081); MACV, *Machupo virus* strain Carvallo (NC_005078); SABV, *Sabiá virus* strain SPH 114202 (NC_006317); TCRV, *Tacaribe virus* strain TRVL II573 (NC_004293).

The Z genes of BCNV strain AV A0070039, TAMV strain W-10777, and WWAV strain AV 9310135 were similar in length to the Z genes of TCRV strain TRVL II573, the 7 South American arenaviruses included in this study, LASV strain Josiah, and LCMV strain WE (Table 2.3). In pairwise comparisons of the Z protein gene sequences, nucleotide sequence identities among the 3 North American viruses ranged from 68.1% to 75.1%, sequence identities between the 3 North American viruses and PICV strain An 3739 and PIRV strain VAV-488 ranged from 60.4% to 62.8%, and sequence identities between the 3 North American viruses and the 6 other New World arenaviruses included in the Z gene sequence alignment ranged from 44.9% to 53.0% (Table 2.4).

Table 2.3. Structural features of the large genomic segments of 13 arenaviruses included in this study

Species	Strain	Length (nucleotides)[a]			GenBank accession no.
		Z protein ORF	IR	RdRp ORF	
Bear Canyon virus	AV A0070039	285	73	6648	AY924391
Tamiami virus	W-10777	285	78	6663	AY924393
Whitewater Arroyo virus	AV 9310135	285	77	6657	AY924395
Pichindé virus	An 3739	285	55	6570	NC_006439
Pirital virus	VAV-488	285	75	6630	NC_005897
Amapari virus	BeAn 70563	285	90	6579	AY924389
Guanarito virus	INH-95551	285	92	6594	NC_005082
Junin virus	XJ13	282	83	6630	NC_005080
Machupo virus	Carvallo	282	82	6627	NC_005079
Sabiá virus	SPH 114202	300	70	6636	NC_006313
Tacaribe virus	TRL II573	285	82	6630	NC_004292
Lassa virus	Josiah	297	100	6654	NC_004297
Lymphocytic choriomeningitis virus	WE	270	196	6627	AF004519

[a]ORF, open reading frame; IR, intergenic region; RdRp, RNA-dependent RNA polymerase.

The neighbor-joining analysis of the *p* model distances generated from Z gene sequence alignment (all 3 nucleotide positions included in the calculation of the genetic distances) indicated that the 3 North American arenaviral species are monophyletic and phylogenetically more closely related to PICV and PIRV than to the 6 other New World arenaviral species included in the Z gene sequence alignment (Figure 2.3).

The RdRp genes of the 3 North American viruses were similar in length to the RdRp genes of TCRV strain TRVL II573, the 7 South American arenaviruses included in this study, LASV strain Josiah and LCMV strain WE (Table 2.3). In pairwise comparisons of the RdRp gene sequences, nucleotide sequence identities among the 3 North American viruses ranged from 66.5% to 67.9%, sequence identities between the 3

North American viruses and PICV strain An 3739 and PIRV strain VAV-488 ranged from 56.5% to 57.8% and sequence identities between the 3 North American viruses and TCRV strain TRVL II573 and the 5 other New World arenaviruses included in the RdRp gene sequence alignment ranged from 49.7% to 51.4% (Table 2.4).

Table 2.4. Nucleotide sequence identities in the Z gene and RNA-dependent RNA polymerase gene of 11 New World arenaviruses[a]

	Z gene (% nucleotide sequence identity)										
Virus[b]	BCN	TAM	WWA	PIC	PIR	AMA	GTO	JUN	MAC	SAB	TCR
BCNV	--	73.3	75.1	60.4	60.7	48.4	53.0	46.4	48.9	48.1	47.3
TAMV	66.5	--	68.1	62.1	60.7	47.3	53.0	46.7	48.2	50.2	46.2
WWAV	67.9	67.7	--	62.8	62.1	48.4	52.7	44.9	47.1	50.2	49.8
PICV	56.5	56.6	56.8	--	68.1	48.7	46.2	44.2	46.7	45.3	44.4
PIRV	57.1	57.8	57.4	60.9	--	47.7	50.5	48.6	50.0	49.5	46.2
AMAV	50.2	50.3	50.5	51.0	51.7	--	64.2	57.8	61.0	54.0	61.1
GTOV	51.1	51.0	51.2	51.2	51.6	67.2	--	64.2	62.8	60.7	64.9
JUNV	50.4	49.7	49.8	51.0	50.8	60.6	61.1	--	68.1	55.3	65.6
MACV	51.4	50.8	50.6	51.3	51.4	61.0	61.4	68.8	--	60.6	67.7
SABV	51.0	51.2	50.8	50.4	51.0	61.2	60.7	61.9	61.6	--	58.2
TCRV	50.5	50.5	50.1	50.9	51.0	60.4	60.6	66.9	67.3	61.6	--
	RNA-dependent RNA polymerase gene (% nucleotide sequence identity)										

[a]Sequence identities among the Z genes and among the RdRp genes are listed above and below the diagonal, respectively.
[b]BCNV, *Bear Canyon virus* strain AV A0070039 (GenBank accession no. AY924391); TAMV, *Tamiami virus* strain W-10777 (AY924393); WWAV, *Whitewater Arroyo virus* strain AV 9310135 (AY924395); PICV, *Pichindé virus* strain An 3739 (NC_006439); PIRV, *Pirital virus* strain VAV-488 (NC_005897); AMAV, *Amapari virus* strain BeAn 70563 (AY924389); GTOV, *Guanarito virus* strain INH-95551 (NC_005082); JUNV, *Junin virus* strain XJ13 (NC_005080); MACV, *Machupo virus* strain Carvallo (NC_005079); SABV, *Sabiá virus* strain SPH 114202 (NC_006313); TCRV, *Tacaribe virus* strain TRVL II573 (NC_004292).

The neighbor-joining analysis of the *p* model distances generated from the RdRp gene sequence alignment (all 3 nucleotide positions included in the calculation of the genetic distances) indicated that the 3 North American arenaviral species are monophyletic and phylogenetically more closely related to PICV and PIRV than to the 6 other New World arenaviral species included in the RdRp gene sequence alignment (Figure 2.3).

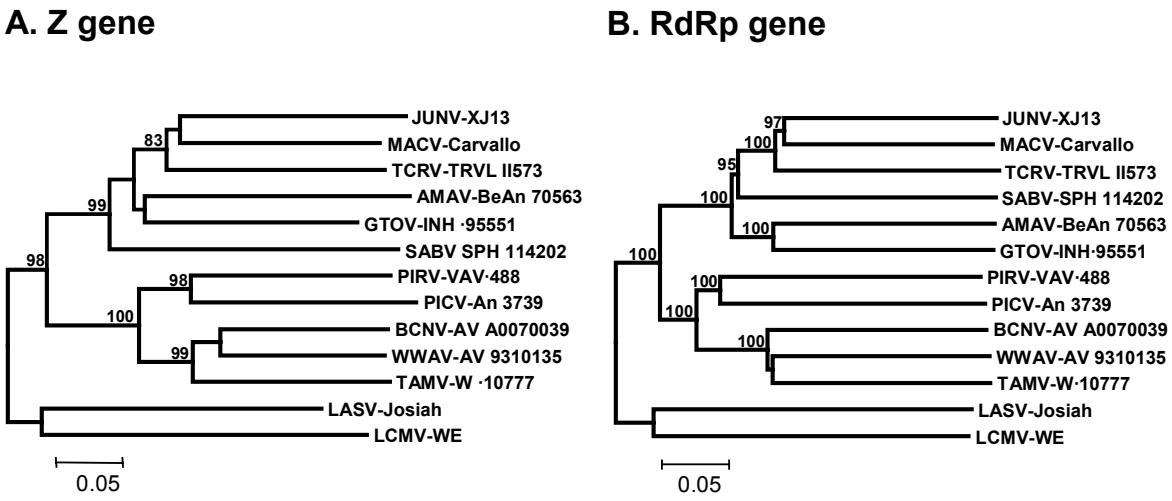


Figure 2.3. Phylogenetic relationships among 13 arenaviruses based on neighbor-joining (NJ) analyses of (A) complete Z gene sequences and (B) complete RNA-dependent RNA polymerase gene sequences. Horizontal branch lengths are proportional to nucleotide sequence divergence. The scale bar represents divergence of 0.05. Bootstrap support was based on 1000 pseudoreplicates; values greater than 50% are shown on each interior branch. BCNV, *Bear Canyon virus* strain AV A0070039 (GenBank accession no. AY924391); TAMV, *Tamiami virus* strain W-10777 (AY924393); WWAV, *Whitewater Arroyo virus* strain AV 9310135 (AY924395); PICV, *Pichindé virus* strain An 3739 (NC_006439); PIRV, *Pirital virus* strain VAV-488 (NC_005897); AMAV, *Amapari virus* strain BeAn 70563 (AY924389); GTOV, *Guanarito virus* strain INH-95551 (NC_005082); JUNV, *Junin virus* strain XJ13 (NC_005080); MACV, *Machupo virus* strain Carvallo (NC_005079); SABV, *Sabiá virus* strain SPH 114202 (NC_006313); TCRV, *Tacaribe virus* strain TRVL II573 (NC_004292).

DISCUSSION

The results of a previous study (Bowen *et al*., 1997) indicated that the principal host relationships of some New World arenaviral species represent a long-term shared

evolutionary relationship between the *Arenaviridae* and the Sigmodontinae. The evidence for this long-standing virus-host relationship includes the present-day association of phylogenetically closely related arenaviral species with phylogenetically closely related sigmodontine rodent species, for example -- JUNV with *Calomys musculinus* (drylands vesper mouse) in Argentina and MACV with *Calomys callosus* (large vesper mouse) in Bolivia (Childs and Peters, 1993). The present-day principal host relationships of other sigmodontine rodent-associated arenaviral species may be a consequence of "host-jumping" or "host-switching" events that involved different sigmodontine rodents or neotomine rodents and sigmodontine rodents.

The isolation of strain AV 98470029 from a large-eared woodrat (*N. macrotis*) in this study is the first direct evidence that BCNV is naturally associated with a rodent species other than the California mouse (*P. californicus*). Other direct evidence that BCNV is naturally associated with the large-eared woodrat includes the isolation of BCNV strain AV B0300052 (GenBank accession no. EF089388) from a large-eared woodrat captured in January 2000 in Zuma Canyon in the Santa Monica Mountains in Los Angeles County, California (M. P. Rood, personal communication). Zuma Canyon is located approximately 140 km northwest of Bear Canyon. The isolation of BCNV strains AV 98470029 and AV B0300052 from large-eared woodrats captured at different sites in southern California suggest that the large-eared woodrat, like the California mouse, is a principal host of BCNV.

The hallmark of the arenaviruses is their ability to establish chronic infections in their principal rodent hosts. The long-term persistence of an arenavirus in nature is dependent upon the capacity of chronically infected rodents to transmit their infections to subsequent generations of conspecifics. Accordingly, future research on the ecology of BCNV should include studies on the duration of BCNV infection in naturally infected large-eared woodrats and California mice and studies on the ability of naturally infected large-eared woodrats and California mice to initiate intraspecific virus transmission.

Collectively, the results of the neighbor-joining analysis of the Z protein, RdRp and N protein gene sequence data in this study support the notion that the L segments and the 3' halves of the S segments of the North American arenaviral species are descended from an ancestor of the South American lineage A viruses that emerged after the last common ancestor of the South American lineage A viruses and the South American lineage B viruses. In contrast, the results of the neighbor-joining analysis of the GPC gene sequence data in this study, as in previous studies (Archer and Rico-Hesse, 2002; Charrel *et al*., 2001, 2002), suggest that the 5' halves of the S segments of the North American arenaviral species are descended from an ancestor of the South American lineage B viruses that emerged after the last common ancestor of the South American lineage A viruses and the South American lineage B viruses. The difference in the placement of the North American lineage with respect to the South American lineages in the phylograms generated from the GPC gene sequence data and from the N protein gene sequence data was attributed in previous studies (Archer and Rico-Hesse, 2002; Charrel *et al*., 2001, 2002) to homologous recombination between the S segment of a lineage A virus and the S segment of a lineage B virus. An alternative explanation for the difference in the placement of the North American lineage with respect to the South American lineages is that the evolution of the GPC genes of the North American arenaviruses has been affected by selection pressures more similar to the selection pressures that shaped the evolution of the GPC genes of the South American lineage B viruses than to the selection pressures that shaped the evolution of the GPC genes of the South American lineage A viruses. Note that this alternative explanation does not entail genetic recombination between a South American lineage A virus and a South American lineage B virus.

The known geographical range of the arenaviruses native to the New World extends from southern California (Fulhorst *et al*., 2002a), southern Utah (Fulhorst *et al*., 2001) and southern Florida (Calisher *et al*., 1970) southward to eastern Bolivia and southern Brazil (Childs and Peters, 1993). The rodent subfamily Neotominae is exclusively North American and includes the white-throated woodrat, other woodrats (*Neotoma* species) that are naturally associated with Tacaribe serocomplex viruses

(Fulhorst *et al*., 2001) and the California mouse. The rodent subfamily Sigmodontinae includes the hispid cotton rat in North America and 11 rodents that serve as principal hosts of arenaviruses native to South America. These rodents are the bicolored arboreal oryzomys (*Oecomys bicolor*) in Peru (ALLV), Guiana bristly mouse (*Neacomys guianae*) in Brazil (AMAV), large-headed oryzomys (*Oryzomys capito*) in Brazil (CPXV), an oryzomys (*Oryzomys* sp.) in Brazil (FLEV), short-tailed cane mouse (*Zygodontomys brevicauda*) in Venezuela (GTOV), drylands vesper mouse (*C. musculinus*) in Argentina (JUNV), large vesper mouse (*C. callosus*) in Bolivia (LATV and MACV), dark bolo mouse (*Bolomys obscurus*) in Argentina (OLVV), Paraguayan oryzomys (*Oryzomys buccinatus*) in Paraguay (PARV), Tomes's oryzomys (*Oryzomys albigularis*) in Colombia (PICV), Alston's cotton rat (*S. alstoni*) in Venezuela (PIRV). Note that *Sigmodon* is the only genus that includes the principal host of a North American arenavirus (TAMV) and the principal host of a South American lineage A virus (PIRV).

The present-day geographical range of the genus *Neotoma* extends from British Columbia southward to Nicaragua and from peninsular Florida westward to coastal California (Musser and Carleton, 2005). In a previous study (Fulhorst *et al*., 2001), arenaviruses phylogenetically closely related to WWAV were isolated from white-throated woodrats (*N. albigula*) captured in northwestern New Mexico and western Oklahoma, a bushy-tailed woodrat (*Neotoma cinerea*) captured in southern Utah, Mexican woodrats (*Neotoma mexicana*) captured in central New Mexico and southern Utah, and southern plains woodrats (*Neotoma micropus*) captured in southern Texas. *Neotoma albigula, N. cinerea, N. mexicana* and *N. micropus* represent 3 of the 4 major phylogenetic subdivisions in the genus *Neotoma* (Edwards and Bradley, 2002). The broad geographical distribution of WWAV and WWA-like viruses in association with distantly related *Neotoma* species indicates that the present-day association between the North American lineage in the *Arenaviridae* and the genus *Neotoma* was established long ago.

Collectively, the isolation of BCNV from large-eared woodrats captured in different counties in southern California, the close phylogenetic relationship between

BCNV and WWAV (and between BCNV and other arenaviruses naturally associated with woodrats in the western United States) and the "ancient" relationship between the *Arenaviridae* and the genus *Neotoma* suggest that the present-day association of BCNV with the large-eared woodrat was established long ago. Hypothetically, the present-day association of BCNV with the California mouse in southern California represents a successful "host-jumping" event from the large-eared woodrat to the California mouse.

The divergence of the sigmodontine rodents from the neotomine rodents has been dated to 16.8 to 18.1 million years ago (Steppan *et al*., 1994). The available fossil record suggests that sigmodontine rodents originally invaded South America from North America during the Pliocene (3.5 to 2.5 million years ago). Thus, the South American arenaviruses likely are descended from an arenavirus (or arenaviruses) that originated in North America.

The present-day association of PIRV with Alston's cotton rat in western Venezuela, ALLV with the bicolored arboreal oryzomys in Peru, FLEV with an oryzomyine rodent in Brazil, PICV with Tomes's oryzomys in Colombia and PARV with the Paraguayan oryzomys in Paraguay suggests that the relationship between the South American lineage A in the *Arenaviridae* and the Sigmodontinae was established long ago -- maybe as early as the divergence of the sigmodontine rodents from the neotomine rodents. Assuming that the association of PIRV with Alston's cotton rat represents a long-term shared evolutionary relationship between the South American lineage A and the Sigmodontinae, then the present-day association of TAMV with *S. hispidus* in southern Florida likely was established after the hispid cotton rat diverged from Alston's cotton rat. Perhaps the present-day association of TAMV with the hispid cotton rat represents a successful host-switching or host-jumping event from a neotomine rodent such as the eastern woodrat (*Neotoma floridana*) to the hispid cotton rat in the eastern United States.

Hypothetically, a successful host-jumping event could lead to an increase in the variety of habitats occupied by infected rodents if the habitat type(s) preferred by the new principal host are not the same as the habitat type(s) preferred by the original principal host. A successful host-jumping event also could lead to an expansion of the geographical range of infected rodents if the new principal host occurs outside the geographical region occupied by the original principal host. As such, a successful host-jumping event may result in a change in the epidemiology of human disease caused by a particular arenavirus.

Chapter 3: Genetic diversity among the South American hemorrhagic fever arenaviruses

INTRODUCTION

The Tacaribe serocomplex viruses native to South America include *Junin virus* (JUNV) and *Oliveros virus* (OLVV) in Argentina, *Machupo virus* (MACV) and *Latino virus* (LATV) in Bolivia, *Amapari virus* (AMAV), *Cupixi virus* (CPXV), *Flexal virus* (FLEV), and *Sabiá virus* (SABV) in Brazil, *Pichindé virus* (PICV) in Colombia, *Parana virus* (PARV) in Paraguay, *Allpahuayo virus* (ALLV) in Peru, and *Guanarito virus* (GTOV) and *Pirital virus* (PIRV) in Venezuela. Four Tacaribe serocomplex viruses naturally cause severe disease in humans in South America (Peters, 2002). Junin virus, MACV, and GTOV are the etiologic agents of Argentine hemorrhagic fever (AHF), Bolivian hemorrhagic fever (BHF), and Venezuelan hemorrhagic fever (VHF), respectively. These diseases are endemic in rural areas of the countries for which they were named (see Figure 3.1). Sabiá virus was the etiological agent in a fatal case of hemorrhagic fever in Sao Paulo, Brazil (Lisieux *et al*., 1994). Infection with the South American hemorrhagic fever arenaviruses is associated with high case-fatality rates (15 - 30%).

Specific members of the subfamily Sigmodontinae in the rodent family Cricetidae (Musser and Carleton, 2005) are the principal hosts of the South American Tacaribe serocomplex viruses for which natural host relationships have been well characterized (Childs and Peters, 1993). For example, the drylands vesper mouse (*Calomys musculinus*) in north-central Argentina is the principal host of JUNV (Mills *et al*., 1992), the large vesper mouse (*C. callosus*) in northeastern Bolivia is the principal host of MACV (Johnson *et al*., 1966), and the short-tailed cane mouse (*Zygodontomys brevicauda*) in western Venezuela is the principal host of GTOV (Fulhorst *et al*., 1999). The natural reservoir of SABV has not been identified. Humans usually become infected with arenaviruses via inhalation of virus in aerosolized droplets of secreta and excreta from

infected rodents or by direct contact with virus through mucosal membranes or percutaneous injuries.

Diagnosis of the 4 South American hemorrhagic fevers usually is based on clinical examination of patients and clinical laboratory findings in conjunction with history of travel to known endemic areas or exposure to rodents. Clinical diagnosis of these arenaviral infections is not definitive because the presentation of these diseases is similar to many other febrile diseases caused by viruses other than arenaviruses, bacteria, and eukaryotic parasites. Laboratory diagnostic assays, most commonly serological assays, are used to confirm the etiology of disease and may include enzyme-linked immunosorbent assay (ELISA), IFA, or neutralization of infectivity (de Manzione *et al.*, 1998; Peters *et al.*, 1973; Riera *et al.*, 1997; Salas *et al.*, 1991). Seroconversion usually occurs late in the course of illness or early convalescence making these tests unreliable for early detection. The gold standard diagnostic test is virus isolation; virus can be isolated from samples of serum, blood, solid tissues, and urine. However, work with hemorrhagic fever arenaviruses requires the use of a biosafety level 4 (BSL-4) laboratory (BMBL, 2007) and work can only be performed by trained personnel. Hence, this diagnostic assay has only been performed in the few BSL-4 laboratories in the world. Results of virus isolation typically are obtained between 7 to 14 days; this test requires considerably more time than serological assays. The duration of the acute illness in the South American arenaviral hemorrhagic fevers is typically 10 to 15 days after onset (Harrison *et al.*, 1999; Jay *et al.*, 2005; Maiztegui, 1975). The results of virus isolation and some of the serological tests may not be known until after the patient has reached convalescence or has progressed to a severe illness beyond benefit of treatment.

Limited therapies exist for the treatment of arenaviral hemorrhagic fevers. Administration of immune plasma for the treatment of AHF can dramatically reduce the mortality from 30% to less than 1% in hospitalized cases (Maiztegui *et al.*, 1979). Passive transfer of immunoglobulin was used in BHF only in the treatment of 4 laboratory-confirmed cases (Stinebaugh *et al.*, 1966). The antiviral drug ribavirin also has been used

to a limited extent in the treatment of AHF and BHF (Enria and Maiztegui, 1994; Kilgore *et al*., 1997). The results of previous studies with AHF cases suggest that the benefit of treatment with immune plasma or ribavirin is dependent upon the length of time between the onset of illness and initiation of therapy (Enria *et al*., 1986; Enria and Maiztegui, 1994; Maiztegui, 1979). Hence, for these therapies to be effective, a rapid and accurate diagnostic test is needed for the detection of infection in acute-phase clinical specimens.

As mentioned previously, most of the serological assays are unreliable in the detection of arenaviruses in acute-phase clinical specimens. The other challenge in using serology-based methods is the lack of specificity of most tests; arenaviruses can be highly cross-reactive in IFA and ELISA. Neutralization is the only test that is highly specific. However, arenavirus-specific antibodies usually do not develop until late in the course of infection. Testing a serum against a number of viruses also requires growth and propagation of infectious virus; work that must be done at BSL-4.

The results of previous studies demonstrated that RT-PCR detection of arenavirus-specific RNA in clinical samples can be used for early diagnosis of arenaviral infections (Bausch *et al*., 2000; Demby *et al*., 1994; Lozano *et al*., 1993, 1995). Sequence-based assays such as RT-PCR assays compared to virus isolation are rapid and can be done without the use of a containment laboratory. Compared to most serological methods, sequence-based methods can be high sensitive and specific. Detection of arenaviral RNA by RT-PCR in acute-phase clinical specimens has only been evaluated for research purposes, in studies on Lassa fever in western Africa (Bausch *et al*., 2000; Demby *et al*., 1994) and AHF (Lozano *et al*., 1995). This diagnostic test is yet to be developed and evaluated for the other South American hemorrhagic fever arenaviruses.

The primary objective of the study described in this chapter was to extend our knowledge of the genetic diversity among and within the arenaviral species that cause hemorrhagic fevers in South America. The results of previous studies using complete GPC gene and N protein gene sequences revealed extensive genetic diversity between

JUNV, MACV, GTOV, and SABV (Archer and Rico-Hesse, 2002; Charrel *et al*., 2002). The results of earlier studies using partial N protein gene sequences or GPC gene sequences demonstrated that there is substantial nucleotide sequence differences among strains of JUNV (Garcia *et al*., 2000) and among strains of GTOV (Weaver *et al*., 2000). The genetic heterogeneity between strains of different arenaviral species (between-species) and among different strains of an arenaviral species (within-species) potentially can pose challenges to the development of rapid, accurate (sensitive and specific) nucleotide sequence-based assays for the detection of arenavirus-specific RNA in acute-phase clinical specimens. Mismatches between oligonucleotide primers used in an RT-PCR assay and the target sequences due to the variability of arenaviral genomes can lower the accuracy of this diagnostic test. Knowledge of the complete genomic sequences of multiple strains of different arenaviral species, between-species genetic diversity, and within-species genetic diversity will be useful in the design of rapid, accurate sequence-based assays for diagnosis of arenaviral infections.

MATERIALS AND METHODS

Viruses

The arenaviral isolates that were genetically characterized in this study included 10 strains of JUNV, 6 strains of MACV, and 7 strains of GTOV. All 23 strains were isolated from laboratory-confirmed AHF, BHF, or VHF cases (see Table 3.1) and were selected from the virus collection of the Special Pathogens Branch, Centers for Disease Control and Prevention (CDC, Atlanta, Georgia).

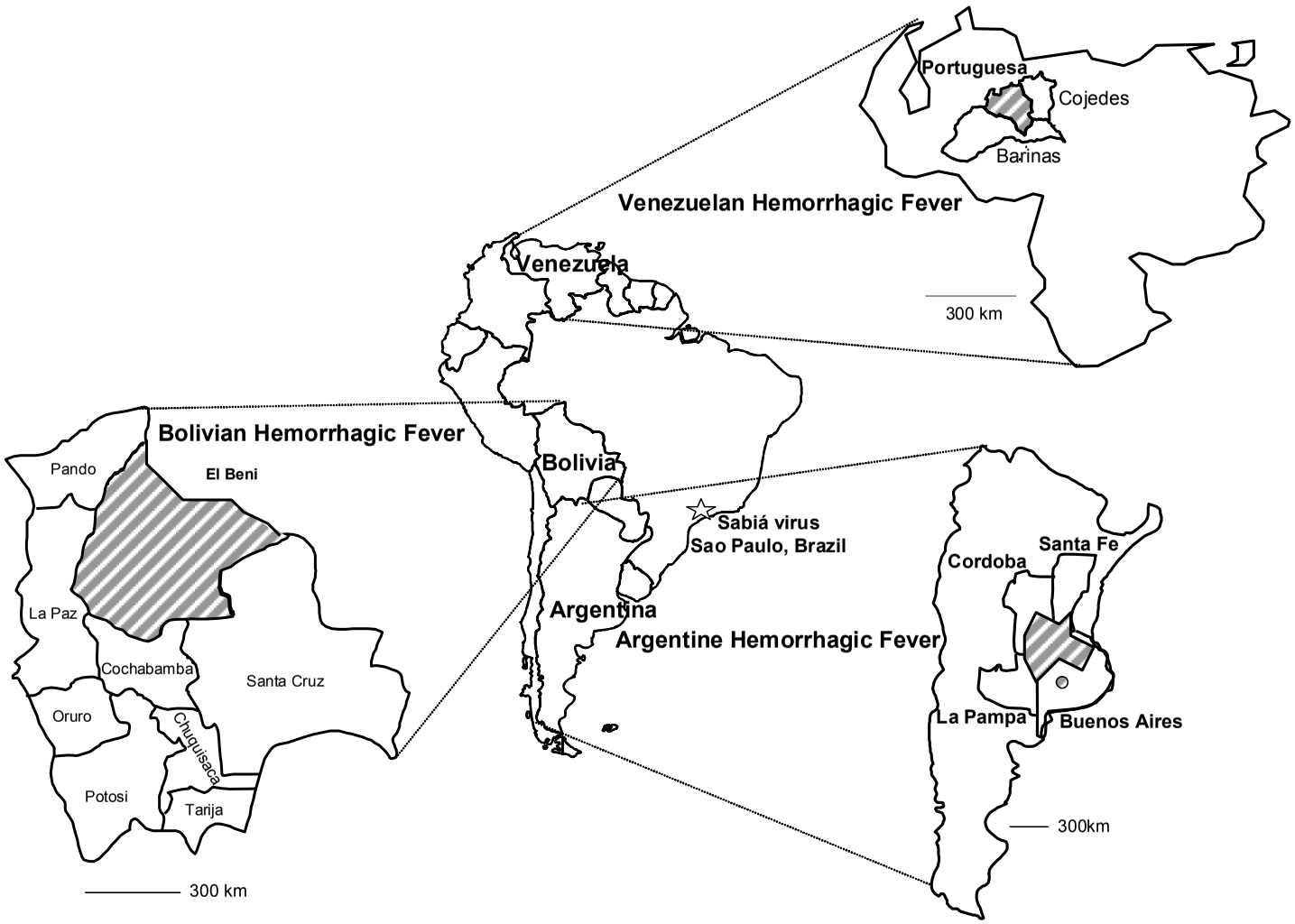


Figure 3.1 Regions in South America endemic for arenaviral hemorrhagic fevers.

Junin virus strains P1998, P2031, P2045, P2290, P3096, P3766, P3778, P3790, P4096, and P35302 were selected to represent different epidemic years of AHF. *Machupo virus* strains MARU216606, MARU249121, MARU222688, 9430084, Chicava, and 200002427 were chosen to represent different genotypes of MACV as determined from a preliminary neighbor-joining analysis of partial N protein gene sequences (data not shown). *Guanarito virus* strains CVH-950801, CVH-960101, CVH-960102, CVH-960103, CVH-960203, CVH-961104, and S-56764 were isolated from VHF cases from different localities in the state of Portuguesa. Other arenaviruses included in the analsyes of between- and within-species genetic diversity included JUNV strains XJ13 and Romero (P3235), MACV strains Carvallo, Mallele (MARU258667), and 9530537, GTOV strain INH-95551, and SABV strain SPH114202 (Table 3.1).

Virus assay and preparation of RNA

Monolayer cultures of Vero E6 cells in 12.5- or 25.0-cm^2 cell culture flasks were inoculated with virus in sterile phosphate buffered saline as described previously (Fulhorst *et al*., 2001). At 7 or 10 days post-inoculation, the virus-infected cell monolayers were lysed in TRIzol® Reagent (Invitrogen Life Technologies, Inc.) or TRI® Reagent BD (Sigma-Aldrich, MO). Propagation of infectious GTOV, MACV, and JUNV was performed by Ms. Mary Louise Milazzo, Dr. Charles F. Fulhorst, and Dr. Pierre E. Rollin inside the Maximum Containment Labortory (a BSL-4 facility) at the CDC. The cell lysates were shipped on dry ice to UTMB. Total RNA was isolated from cell lysates in the BSL-2 laboratory at UTMB.

Table 3.1 Histories of 30 viruses isolated from cases of arenaviral hemorrhagic fever

Species[a]	Strain	Isolated from			GenBank Accession no.
		Locality[b]	Country	Year[c]	S segment
JUNV	P1998	Buenos Aires	Argentina	1972	DQ854730
JUNV	P2031	Buenos Aires	Argentina	1972	DQ854731
JUNV	P2045	Buenos Aires	Argentina	1972	DQ854733
JUNV	P2290	Buenos Aires	Argentina	1973	DQ854736
JUNV	P3096	Buenos Aires	Argentina	1975	DQ854737
JUNV	P3778	Buenos Aires	Argentina	1977	DQ854734
JUNV	P3766	Buenos Aires	Argentina	1977	DQ854735
JUNV	P3790	Buenos Aires	Argentina	1977	DQ854739
JUNV	P4036	Buenos Aires	Argentina	1979	DQ854732
JUNV	P35302	Buenos Aires	Argentina	1991	DQ854738
MACV	MARU 216606	El Beni	Bolivia	1964	AY924206
MACV	MARU 249121	El Beni	Bolivia	1964	AY924208
MACV	MARU 222688	El Beni	Bolivia	1969	AY924207
MACV	Chicava	El Beni	Bolivia	1993	AY924202
MACV	9430084	El Beni	Bolivia	1994	AY924203
MACV	200002427	El Beni	Bolivia	2000	AY924204
GTOV	CVH-950801	Portuguesa	Venezuela	1995	AY572557
GTOV	CVH-960101	Portuguesa	Venezuela	1996	AY497548
GTOV	CVH-960102	Portuguesa	Venezuela	1996	AY572558
GTOV	CVH-960103	Portuguesa	Venezuela	1996	AY572556
GTOV	CVH-960302	Portuguesa	Venezuela	1996	AY572555
GTOV	CVH-961104	Portuguesa	Venezuela	1996	AY572561
GTOV	S-56764	Portuguesa	Venezuela	1997	AY572554
JUNV	XJ13	Buenos Aires	Argentina	1958	NC_005081[d]
JUNV	Romero (P3235)	Buenos Aires	Argentina	1976	AY619641[d]
MACV	Carvallo	El Beni	Bolivia	1963	AY129248[d]
MACV	Mallele (MARU258667)	Not known	Bolivia	1971	AY619645[d]
MACV	9530537	La Paz	Bolivia	1995	AY571959[d]
GTOV	INH-95551	Portuguesa	Venezuela	1990	NC_005077[d]
SABV	SPH114202	Sao Paulo	Brazil	1990	NC_006317[d]

[a] JUNV, *Junin virus;* MACV, *Machupo virus*; GTOV, *Guanarito virus*; SABV, *Sabiá virus*.
[b]Province in Argentina; department in Bolivia; state in Venezuela. Each province/state/department consists of smaller units known as towns, municipalities, or localities. Each of the GTOV strains was isolated from different localities in the state of Portuguesa.
[c]Year in which the specimen was collected.
[d]The S genomic segment sequences were previously determined and made available in the GenBank sequence database.

Synthesis of first-strand cDNA, PCR amplification, and sequencing

Reverse transcription was carried out on the total RNA following the protocol described previously (Günther *et al*., 2000; Cajimat and Fulhorst, 2004) by using SuperScript II RNase H^- Reverse Transcriptase or SuperScript III RNase H^- Reverse Transcriptase (Invitrogen Life Technologies, Inc.) in conjunction with oligonucleotide 19C-cons. As described previously, oligonucleotide 19C-cons is a derivative of oligonucleotide ARE3'END (Gonzalez *et al*., 1995) and was expected to anneal to the 19-nt fragment at the extreme 3' end of the S segment genomic RNA and replicative intermediate (Auperin *et al*., 1982a,b; Southern, 1996).

The nucleotide sequences of the S genomic segments of the 23 South American arenaviruses were determined from 3 overlapping fragments designated S1, S2, and S3 (Figure 3.2), similar to the technique described previously by Charrel and others (2001). S1 extends from within the 5' NCR to the 3'end of the GPC gene; S2 extends from within the 3'end (~200 nt) of the GPC gene, the intergenic region (IR), to the 5'end (~200 nt) of the N protein gene; S3 extends from within the 5'end of the N protein gene to the 3' NCR. Oligonucleotide primers used to amplify the 3 fragments were initially designed based on the S segment sequences of JUNV strains XJ13 (GenBank Accession no. NC_005081) and MC2 (D10072), MACV strain Carvallo (AY129248), GTOV strain INH-95551 (NC_005077), AMAV strain BeAn 70563 (AF512834), CPXV strain BeAn119303 (AF512832), and SABV strain SPH 114202 (NC_006317). Initial primers used for PCR amplification and sequencing were sensitive to multiple strains. Additional oligonucleotide primers for sequencing were designed as sequence data from multiple strains became available.

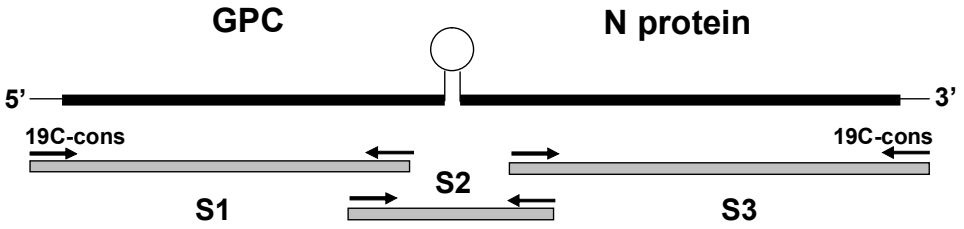


Figure 3.2. **Strategy for PCR amplification and sequencing of the S genomic segments of 23 South American arenaviruses**. The strategy was similar to that described by Charrel and others (2001). The GPC gene and the N protein gene are represented by thick solid lines; the intergenic region (hairpin) and noncoding region at the 5' and 3' ends are represented by thin lines; the 3 fragments S1, S2, and S3, are represented by grey-colored lines. Regions where oligonucleotide primers anneal are represented by small arrows. Oligonucleotide primers and their sequences are described in the text.

The S1 fragment was generated from the first strand cDNA by using Master Taq Kit (Eppendorf North America) in conjunction with oligonucleotide primer 19C-cons *and* species-specific GPC gene/IR oligonucleotide primers: JUNV - AVGPC16 (5'-CCACAYCTRCAASCWCC-3') or AVGPC97 (5'-ATGGGGCAGACAATCAATGC-3'); MACV - AVGPC16 or AVGPC36 (5'-GCCCGGGTCGGCAGGGGTCTTG-3'); GTOV - AVGPC16, AVGPC35 (5'-GCCGACAGAGCTGGAGGGCTG-3'), or AVNP37 (5'-TGCCYYCCCAGTCCGCGG-3'). The PCR parameters consisted of initial denaturation at 94°C for 2 minutes, followed by 35 cycles of 94°C (35 seconds), 50°C (1 minute), 72°C (1.5 minutes), and a final extension at 72°C of 5 minutes. Similarly, the S3 fragment was amplified using 19C-cons *and* species-specific N protein oligonucleotide primers: JUNV - AVNP122 (5'-GCCGCGGACTGGGGAGGCACTG-3'); MACV - AVNP42 (5'-GCCGCGGACTGGGAGGGCA-3'); GTOV - AVNP42. The same PCR parameters as the amplification of fragment S1 were used. The S2 fragment was amplified by using primers that were designed based on the sequences obtained from both S1 and S3 fragments. The S2 fragment of JUNV was amplified using oligonucleotide primers AVGPC73 (5'-TGGAGCACAGTATTCTTCACAGC-3') *and* AVNP99 (5'-CTTACATCTGAACAAGCCAGGCA-3'); the S2 fragment of MACV was amplified using oligonucleotide primers AVGPC52 (5'-AGTGATCACCTCATC-

TCAGAGAT-3') *and* AVNP71 (5'-TCAAGCAAGACTCTATGATCA-GC-3') or AVNP85 (5'-TCCAGTGAACCTCAYTGTGCTC-3') or AVNP86 (5'-TCTGGTGAA-CCTCAYTGTGCTC-3'); the S2 fragment of GTOV was amplified using oligonucleotide primers AVGPC26 (5'-CCGTGCCCTTTGCCTCACAGGC-3') *and* AVNP52 (5'-CAGTCTGTTCTTGATGGTCATC-3') or AVNP53 (5'- TGTTCCDAAGAAGAA-AAACAAGG-3').The S2 fragment was amplified at a higher annealing of 55°C, presumably to melt the secondary structure present in the IR.

The PCR products generated from the 3 fragments of the first-strand S segment cDNA were purified by using the QIAquick gel extraction kit (Qiagen, Inc.) or QIAEX II gel extraction kit (Qiagen, Inc.). Both strands of each amplicon were sequenced directly by using the dye termination cycle sequencing technique (Applied Biosystems, Inc.). The S2 fragments were sequenced in the presence of DMSO or SequenceR$_x$ Enhancer Solution A (Invitrogen Life Technologies, Inc.). The nucleotide sequences of the S segments of the 23 arenaviruses were deposited to the GenBank sequence database (see Table 3.1).

Data analyses

The nucleotide sequences of the GPC genes and N protein genes of the 23 viruses characterized in this study were compared with the homologous sequences of 7 viruses: JUNV strains XJ13 (GenBank Accession no. NC_005081) and Romero (AY619641); MACV strains Carvallo (AY129248), Mallele (AY619645), and 9530537 (AY571959); GTOV strain INH-95551 (NC_005077). The predicted amino acid sequences of the GPC gene and the N protein gene of all 30 viruses were aligned independently by using the computer program ClustalW 1.7 (Thompson *et al*., 1994). The corresponding nucleotide sequences were aligned by using ClustalW 1.83 in the computer software package MEGA version 3.1 (Kumar *et al*., 2004). The uncorrected *p*-model distances, all 3 nucleotide positions included in the distance calculations, were calculated from the

multiple nucleotide sequence alignments by using programs in the computer software package MEGA version 3.1 (Kumar *et al*., 2004).

RESULTS

The lengths of the nucleotide sequences of the S genomic segments of the 10 strains of JUNV characterized in this chapter, JUNV strain XJ13, and JUNV strain Romero ranged from 3408-nt to 3411-nt; the lengths of the S segments of the 6 strains of MACV characterized in this chapter and MACV strains Carvallo, Mallele (MARU258667), and 9530537 were identical, 3439-nt; and the lengths of the S segments of the 7 strains of GTOV characterized in this study and GTOV strain INH-95551 ranged from 3342- to 3343-nt. Each S genomic segment sequence extends from the 5' NCR (including the terminal 19-nt used as oligonucleotide primer), the complete GPC gene, IR, complete N protein gene, and the 3' NCR (including the terminal 19-nt used as oligonucleotide primer). The structural features of the S genomic segments of the 30 viruses are listed in Table 3.2. The lengths of the GPC gene and the N protein gene were conserved among the different strains of a species.

Between-species genetic diversity among the South American hemorrhagic fever arenaviruses

The GPC and N protein amino acid sequence alignments were 563 and 579 characters in length, respectively. The corresponding GPC gene and N protein gene nucleotide sequence alignments were 1689 and 1737 characters, respectively. Each multiple sequence alignment consisted of the complete gene sequences of 30 viruses: 12 strains of JUNV, 9 strains of MACV, 8 strains of GTOV, 1 strain of SABV.

Table 3.2 Structural features of the small (S) genomic segments of 30 South American arenaviruses[a]

Species	Strain	GenBank Accession no.	Length (nucleotides)[b]			
			GPC gene	IR	N protein gene	S segment
JUNV	P1998	DQ854730	1455	91	1692	3409
JUNV	P2031	DQ854731	1455	91	1692	3409
JUNV	P2045	DQ854733	1455	91	1692	3409
JUNV	P2290	DQ854736	1455	91	1692	3410
JUNV	P3096	DQ854737	1455	91	1692	3409
JUNV	P3766	DQ854735	1455	91	1692	3409
JUNV	P3778	DQ854734	1455	91	1692	3408
JUNV	P3790	DQ854739	1455	91	1692	3409
JUNV	P4036	DQ854732	1455	91	1692	3409
JUNV	P35302	DQ854738	1455	91	1692	3409
MACV	MARU216606	AY924206	1488	90	1692	3439
MACV	MARU249121	AY924208	1488	90	1692	3439
MACV	MARU222688	AY924207	1488	90	1692	3439
MACV	Chicava	AY924202	1488	90	1692	3439
MACV	9430084	AY924203	1488	90	1692	3439
MACV	200002427	AY924204	1488	90	1692	3439
GTOV	CVH-950801	AY572557	1437	94	1680	3342
GTOV	CVH-960101	AY497548	1437	94	1680	3342
GTOV	CVH-960102	AY572558	1437	94	1680	3342
GTOV	CVH-960103	AY572556	1437	94	1680	3342
GTOV	CVH-960302	AY572555	1437	94	1680	3342
GTOV	CVH-961104	AY572561	1437	94	1680	3342
GTOV	S-56764	AY572554	1437	94	1680	3342
JUNV	XJ13	AY358023[c]	1458	91	1692	3411
JUNV	Romero (P3235)	AY619641[c]	1455	91	1692	3409
MACV	Carvallo	AY129248[c]	1488	90	1692	3439
MACV	Mallele (MARU258667)	AY619645[c]	1488	90	1692	3439
MACV	9530537	AY571959[c]	1488	90	1692	3439
GTOV	INH-95551	NC_005077[c]	1437	95	1680	3343
SABV	SPH114202	NC_006317[c]	1464	94	1686	3366

[a]The S genomic segments of 23 viruses were determined in this chapter and that of 7 strains were published previously or made available in the GenBank sequence database.
[b]Lengths of glycoprotein precursor (GPC) gene and nucleocapsid (N) protein gene include the start codon to the last amino acid of the open reading frame. The length of the intergenic region (IR) includes the region in between the stop codon of the GPC gene and the stop codon of the N protein gene. The length of the S genomic segment includes the 5'NCR upstream of the GPC gene and the 3'NCR downstream of the N protein gene, including the 19-nt fragments at each end that correspond to the sequence of oligonucleotide primer 19C-cons.
[c]The S segment sequences of these viruses were obtained from the GenBank sequence database.

The results of pairwise comparisons of complete gene (nucleotide) sequences between the 30 viruses are summarized in Table 3.3. In pairwise comparisons of complete GPC gene sequences, the 12 strains of JUNV exhibited the least sequence nonidentity with the 9 strains of MACV (30.7% to 32.6%), and the 8 strains of GTOV exhibited the lowest sequence nonidentity with SABV (41.0% to 41.8%). Similarly, in

pairwise comparisons of complete N protein gene sequences, the 12 strains of JUNV exhibited the least sequence nonidentity with the 9 strains of MACV (22.0% to 24.6%), and the 8 strains of GTOV exhibited the lowest sequence nonidentity with MACV (29.6% to 31.4%).

Table 3.3. Nucleotide sequence nonidentities in the glycoprotein precursor gene and nucleocapsid protein gene of 30 viruses[a]

	GPC gene (% nucleotide sequence nonidentity)[b]			
Species	*Junin virus*	*Machupo virus*	*Guanarito virus*	*Sabiá virus*
Junin virus	-	30.7 – 32.6	41.3 – 42.5	42.2 – 43.0
Machupo virus	22.0 – 24.6	-	42.3 – 44.2	41.8 – 43.3
Guanarito virus	30.1 – 31.8	29.6 – 31.4	-	41.0 – 41.8
Sabiá virus	32.7 – 33.6	32.0 – 33.0	33.8 – 34.3	-
	N protein gene (% nucleotide sequence nonidentity)			

[a]The nucleotide sequences of the glycoprotein precursor (GPC) gene and nucleocapsid (N) protein gene of 12 strains of *Junin virus* (JUNV), 9 strains of *Machupo virus* (MACV), 8 strains of *Guanarito virus* (GTOV), and 1 strain of SABV were compared.
[b]Sequence nonidentities in the GPC gene and in the N protein gene are located above and below the diagonal, respectively.

Within-species genetic diversity among the South American hemorrhagic fever arenaviruses

Junin virus

In pairwise comparisons of the nucleotide sequences of the 12 JUNV strains, GPC gene sequence nonidentities and N protein gene sequence nonidentities ranged from 0.1 to 4.9% and from 0.0 to 4.3%, respectively (Table 3.4). The lowest GPC gene sequence nonidentity was observed between strains P3766 and P3790, both isolated in the same year and probably from the same epidemic. The most divergent sequences were obtained between strain P35302 and strain P3778. In the N protein gene, the lowest sequence nonidentity again was observed between strains P3766 and P3790, and the highest sequence nonidentity was observed between strains P2290 and P3778. The mean

percentage sequence nonidentity in the GPC gene and in the N protein gene among the 12 strains of JUNV was 3.1% and 2.8%, respectively.

Table 3.4. Nucleotide sequence nonidentities between strains of arenavirus species[a]

	GPC gene		N protein gene	
Species	range	mean	range	mean
Junin virus	0.1 - 4.9	3.1	0.0 - 4.3	2.8
Machupo virus	0.1 - 13.0	8.4	0.2 - 12.0	7.5
Guanarito virus	0.1 – 7.1	3.6	0.2 – 7.5	3.7

[a] *Junin virus*, 12 strains; *Machupo virus*, 9 strains; *Guanarito virus*, 8 strains; GPC, glycoprotein precursor; N, nucleocapsid.

Machupo virus

In pairwise comparisons of the nucleotide sequences of the 9 MACV strains, GPC gene sequence nonidentities and N protein gene sequence nonidentities ranged from 0.1% to 13.0% and from 0.2% to 12.0%, respectively. The lowest GPC gene sequence nonidentity was between strains Carvallo and MARU216606 whereas the highest GPC gene sequence nonidentity was observed between strains Mallele (MARU258667) and 9530537. Similarly, the lowest N protein gene sequence nonidentity was between strains Carvallo and MARU216606 and the highest N protein gene sequence nonidentity was observed between strains 9530537 and 200002427. The mean sequence nonidentity in the GPC gene and in the N protein gene among the 9 strains of MACV was 8.4% and 7.5%, respectively.

Guanarito virus

In pairwise comparisons of the nucleotide sequences of the 8 GTOV strains, sequence nonidentities in the GPC gene and in the N protein gene ranged from 0.1% to 7.1% and from 0.2% to 7.5%, respectively. The lowest GPC gene sequence nonidentity was observed between strains CVH-960103 and CVH-960302 and the highest GPC gene sequence nonidentity was observed between strains CVH-961104 and S-56764.

Similarly, the lowest N protein gene sequence nonidentity was between strains CVH-960101 and CVH-960102 and the highest N protein gene sequence nonidentity was observed between strains S-56764 and CVH-960101, CVH-960102, or CVH-950801.

DISCUSSION

Four South American arenaviruses naturally cause severe febrile diseases in humans (Peters, 2002). Argentine hemorrhagic fever was described as a clinical entity in the mid-1950's and the etiologic agent, Junin virus, subsequently was isolated from human cases during an outbreak in the province of Buenos Aires (Maiztegui, 1975). The AHF epidemic area encompasses the richest farmlands of Argentina and includes more than half of the total population of the country, making AHF a major public health problem (Maiztegui, 1975). Since the early 1960's, outbreaks of AHF have been documented annually in northern-central Argentina within the provinces of Buenos Aires, Cordoba, La Pampa, and Santa Fe (Maiztegui, 1975). Bolivian hemorrhagic fever was first described in 1959 in the Beni Department of Bolivia (MacKenzie *et al*., 1964). Subsequently, the etiologic agent, Machupo virus, was isolated from 5 cases during an investigation of an ongoing epidemic (Johnson *et al*., 1965). Outbreaks of BHF have occurred sporadically in northeastern Bolivia from the early 1960's to 1994 (CDC, 1994; Kilgore *et al*., 1995; MacKenzie *et al*., 1964; Peters *et al*., 1974). Venezuelan hemorrhagic fever was first recognized as a distinct clinical entity in 1989 (Salas *et al*., 1991). Guanarito virus subsequently was isolated and identified as the etiologic agent of VHF (Tesh *et al*., 1994). The disease is restricted to a small area within the region in which GTOV is enzootic (Fulhorst *et al*., 1999; de Manzione *et al*., 1998; Weaver *et al*., 2000). Sabiá virus was the cause of a single fatal case of hemorrhagic fever in Sao Paulo, Brazil (Lisieux *et al*., 1994). The diseases caused by these 4 arenaviruses are collectively known as the South American arenaviral hemorrhagic fevers.

As mentioned previously, clinical diagnosis of the hemorrhagic fevers caused by JUNV, MACV, GTOV, and SABV is difficult because the early symptoms and

presentation of the diseases caused by these viruses are common with diseases caused by other infectious agents. The most sensitive laboratory diagnostic test to confirm the etiology of disease is virus isolation which typically takes 7 to 14 days; hence, this test is not a rapid diagnostic test. The development of sequence-based assays as substitute for virus isolation and serological assays to detect arenaviral infections is highly desirable. However, this avenue has not been thoroughly investigated for the 4 South American hemorrhagic fever arenaviruses. There has only been one published study utilizing RT-PCR to detect arenavirus-specific RNA in acute-phase blood samples (Lozano *et al*., 1995). The results of this study suggested that detection of JUNV by RT-PCR is equally sensitive if not more sensitive than conventional serology-based methods (Lozano *et al*., 1995). However, the assay described in this study lacked specificity when compared with the results of ELISA and/or neutralization tests (Lozano *et al*., 1995).

Sequence-based assays for arenaviral infections can be highly sensitive and specific (Bausch *et al*., 2000; Demby *et al*., 1994). The sensitivity of RT-PCR assays for detection of arenavirus-specific RNA may be affected by mismatches between the oligonucleotide primers and the target sequences due to the genetic variability of arenaviral genomes. The results of previous studies revealed extensive genetic diversity between JUNV, MACV, GTOV, and SABV (Archer and Rico-Hesse, 2002; Charrel *et al*., 2002). The high genetic diversity among the S segments of these 4 viruses in combination with the lack of complete gene sequence data from multiple strains have hampered the efforts of our laboratory as well as other groups from developing an RT-PCR assay using universal oligonucleotide primers that can detect all 4 arenaviral species (Vieth *et al*., 2005).

The results of previous studies using a short fragment of the GPC gene (~390-nt) or the N protein gene (~616-bp) indicated that there is extensive genetic diversity within each of various arenaviral species (Bowen *et al*., 2000; Garcia *et al*., 2000; Weaver *et al*., 2000, 2001). One reason may be that this region of the N protein gene is more variable than the rest of the S segment; hence it is useful in phylogenetic studies but cannot be

used in assessment of strain-to-strain genetic variation for the purpose of designing RT-PCR based detection assays. The results of the study described in this chapter also reveal that there is a high degree of genetic heterogeneity in the genomes of strains within a species and this genetic variability may pose challenges in the development of sequence-based assays that can be used to detect strains across the geographic range of a species.

Collectively, the nucleotide sequences generated in this study combined with the nucleotide sequences previously published will be useful in the design of sequence-based assays for rapid detection of arenavirus-specific RNA in acute-phase clinical specimens. Knowledge of the genetic diversity between strains of different species could be used to develop assays to detect the 4 South American hemorrhagic fever arenaviruses or assays that distinguish a particular species especially in areas where multiple arenaviruses co-exist. Knowledge of the genetic diversity among the different strains of a single species could help us to develop sequence-based assays that can detect even the most divergent strains of a particular species.

There is no specific therapy for the South American arenaviral hemorrhagic fevers. The results of studies done in Argentina demonstrated that administration of immune plasma early in the course of disease can dramatically reduce the mortality from 30% to less than 1% of hospitalized AHF cases (Maiztegui *et al*., 1979). This mode of treatment was tested in 4 cases of BHF (Stinebaugh *et al*., 1966) but has not been tested in VHF. The ability to neutralize JUNV *in vitro* with monoclonal antibodies specific to the GPC (Sanchez *et al*., 1989) suggests that therapy with monoclonal antibodies may be beneficial in the treatment of the South American arenaviral hemorrhagic fevers. However, it has been documented in AHF that the benefit of treatment with immune plasma is dependent on the length of time between the onset of acute illness and initiation of therapy (Enria *et al*., 1986; Enria and Maiztegui, 1994; Maiztegui, 1979). For passive antibody therapy to be given early in the course of disease, a rapid and accurate diagnostic test is needed for the detection of the 4 hemorrhagic fever arenaviruses.

The success of passive antibody therapy has been positively associated with the capacity of an immune serum to neutralize infectivity *in vitro* and can vary from strain to strain within an arenavirus species (Enria *et al*., 1986). Knowledge of the sequence diversity in the neutralizing epitopes located in the GPC will be useful in assessing the ability of monoclonal antibodies to neutralize divergent strains of a species and in the development of monoclonal antibodies that can specifically neutralize strains with divergent sequences in the GPC.

Chapter 4: Genetic criteria for species demarcation in the family *Arenaviridae*

INTRODUCTION

The *Arenaviridae* comprises 2 serocomplexes and 22 species (Salvato *et al*., 2005). The lymphocytic choriomeningitis-Lassa serocomplex includes *lymphocytic choriomeningitis virus* (LCMV), *Ippy virus* (IPPYV), *Lassa virus* (LASV), *Mobala virus* (MOBV), and *Mopeia virus* (MOPV). The Tacaribe serocomplex includes *Bear Canyon virus* (BCNV), *Tamiami virus* (TAMV), and *Whitewater Arroyo virus* (WWAV) in North America, *Tacaribe virus* (TCRV) on Trinidad, and *Allpahuayo virus* (ALLV), *Amapari virus* (AMAV), *Cupixi virus* (CPXV), *Flexal virus* (FLEV), *Guanarito virus* (GTOV), *Junin virus* (JUNV), *Latino virus* (LATV), *Machupo virus* (MACV), *Oliveros virus* (OLVV), *Parana virus* (PARV), *Pichindé virus* (PICV), *Pirital virus* (PIRV), and *Sabiá virus* (SABV) in South America.

Specific members of the rodent families Muridae and Cricetidae (Musser and Carleton, 2005) are the principal hosts (reservoirs) of the arenavirus species for which natural host relationships have been well studied. For example, the ubiquitous house mouse (*Mus musculus*) is the principal host of LCMV (Childs and Peters, 1993) and the hispid cotton rat (*Sigmodon hispidus*) in southern Florida is the principal host of TAMV (Calisher *et al*., 1970; Jennings *et al*., 1970).

Six arenaviruses are etiological agents of severe disease in humans. Lymphocytic choriomeningitis virus is a cause of aseptic meningitis and congenital abnormalities (Barton *et al*., 1995; Jahrling and Peters, 1992) and has been associated with organ transplant-associated systemic illnesses and deaths (CDC, 2005). Lassa virus is the agent of Lassa fever, an acute febrile illness that is endemic in western Africa (McCormick and Fisher-Hoch, 2002). Guanarito virus, JUNV, MACV, and SABV are the agents of Venezuelan hemorrhagic fever, Argentine hemorrhagic fever, Bolivian hemorrhagic

fever, and hemorrhagic fever in Brazil, respectively. Work with infectious LASV, GTOV, JUNV, MACV, and SABV requires the highest level of biocontainment, i.e., biosafety level (BSL)-4 (BMBL, 2007). There are only a few BSL-4 laboratories in the world.

Historically, complement fixation (CF), indirect fluorescent antibody (IFA), neutralization of infectivity, and other serological methods were used to define the taxonomical relationships among the members of the *Arenaviridae* (Casals *et al*., 1975; Rowe *et al*., 1970b; Webb *et al*., 1969). The same serological methods have been used to identify arenaviruses isolated from human specimens or rodent tissues (Peters, 2002). Strains of different arenavirus species can be highly cross-reactive in CF and IFA tests. Thus, these serological methods sometimes lack the specificity necessary to distinguish between strains of different arenavirus species. In contrast, neutralization of infectivity usually is highly specific to an arenavirus species.

Neutralization assays and preparation of infectious virus stocks and other materials required for these assays require work with infectious virus *in vitro* (cultured cells) or *in vivo* (small laboratory rodents). Thus, neutralization assays to establish the taxonomical relationship of a potentially novel arenavirus with LASV and the South American hemorrhagic fever arenaviruses require access to a BSL-4 laboratory.

The first comprehensive study on the phylogeny of the *Arenaviridae* was done in 1996 and based on nucleotide sequence data generated from a small fragment of the N protein gene (Bowen *et al*., 1996, 1997). The results of phylogenetic analyses using partial N protein gene sequences were highly concordant with the taxonomy based on the results of various serological assays. For example, the 19 arenaviruses that were included in the study were divided phylogenetically into 2 major clades: Old World (LCMV, IPPYV, LASV, MOBV, and MOPV) and New World (TCRV, AMAV, FLEV, GTOV, JUNV, LATV, MACV, OLVV, PARV, PICV, PIRV, SABV, TAMV, and WWAV) (Bowen *et al*., 1997).

Recent studies have utilized full-length GPC and N protein sequences to establish the taxonomical and phylogenetic relationships between potentially novel arenaviruses and strains of the arenavirus species included in the family *Arenaviridae*, as defined by the International Committee on Taxonomy of Viruses (ICTV) (Salvato *et al*., 2005). With the recent discovery of "novel" arenaviruses and the availability of S genomic segment sequences in the GenBank sequence database, the relationships of a new arenavirus to existing arenavirus species have been determined primarily by comparisons of complete GPC gene and N protein gene sequences (Moncayo *et al*., 2001; Charrel *et al*., 2002).

The criteria currently used by the ICTV to define a species in the genus *Arenavirus* are 1) "association with a specific host species or group of species", 2) "presence in a defined geographical area", 3) "etiological agent (or not) of human disease", 4) "significant differences in antigenic cross-reactivity", and 5) "significant amino acid sequence differences" (Salvato *et al*., 2005). The primary objective of the work summarized in this chapter was to provide a measure of a meaningful lower limit for "significant amino acid sequence differences" in the context of species demarcation within the family *Arenaviridae*. This information could be used by taxonomists, including members of the ICTV, in determination of whether a specific arenavirus isolate is a strain of an established species or a novel species within the *Arenaviridae*.

A previous study suggested a lower limit or cutoff value of 12% amino acid sequence nonidentity (uncorrected *p* model distances) in the N protein to distinguish between strains of different arenavirus species (Bowen *et al*., 2000). This cutoff value was based on the results of pairwise comparisons of the predicted amino acid sequences of the GPC and N protein of LASV strains Josiah, GA391, LP, and 803213. This proposed lower limit has only been used in the context of species demarcation on multiple strains of LASV (Bowen *et al*., 2000; Emonet *et al*., 2006); it remains to be evaluated among the New World arenaviruses because of the lack of complete genomic sequences for multiple strains of any of the arenavirus species indigenous to the Americas. In the study by Bowen and others (2000), comparisons of partial N protein

sequences among 57 strains of LASV revealed up to 14.8% sequence nonidentity. It is not known whether the same degree of genetic diversity occurs when complete gene sequences are compared among these strains of LASV. The results of a previous study (Weaver *et al*., 2001) revealed up to 16% sequence nonidentity in comparisons of partial N protein sequences among 30 strains of PIRV. This suggests that the genetic diversity within the New World arenaviral species may exceed what is seen among LASV.

In this chapter, we used the complete GPC and N protein sequences of 42 strains belonging to 4 South American arenavirus species to improve our knowledge of strain-to-strain variation within these New World arenavirus species and genetic diversity between arenavirus species in the *Arenaviridae*. The 42 strains included 12 strains of GTOV, 5 strains of PIRV, 15 strains of JUNV, and 10 strains of MACV. The sequences of 23 strains were determined in the previous chapter, the sequences of 8 were published previously, and the sequences of 11 strains were determined in this chapter. The examination of within-species diversity was extended to include the nucleotide sequences of 7 strains of LASV that were available in the GenBank sequence database. An objective of the study described in this chapter was to determine the genetic identity of the 11 arenavirus strains isolated from rodents: 4 isolates from the short-tailed cane mouse (*Zygodontomys brevicauda*) in Venezuela (strains VHF-1608, VHF-1750, VHF-3990, and AV 97021119), 4 isolates from the Alston's cotton rat (*Sigmodon alstoni*)in Venezuela (strains 1645, 1743, 3945, and AV 97021016), 2 isolates from the yellow pygmy rice rat (*Oligoryzomys flavescens*) in Argentina (strains JNM-6682 and JNM-7354), and 1 isolate from large vesper mouse (*Calomys callosus*) in Bolivia (strain 9301012).

MATERIALS AND METHODS

Genetic characterization of Tacaribe serocomplex viruses

Guanarito virus strains VHF-1608, VHF-1750, and VHF-3990 originally were isolated from short-tailed cane mice (*Z. brevicauda*) captured in western Venezuela during field studies on the epidemiology of Venezuelan hemorrhagic fever in the 1990's (Weaver *et al*., 2000). *Pirital virus* strains 1645, 1743, and 3945 originally were isolated from Alston's cotton rats (*S. alstoni*) captured in western Venezuela during the field studies on the epidemiology of Venezuelan hemorrhagic fever in the 1990's (Weaver *et al*., 2001). Strains AV 97021119 and AV 97021016 were isolated from a cane mouse and cotton rat, respectively, during a field study on the ecology of hantaviruses enzootic in western Venezuela (C. F. Fulhorst, unpublished data). Arenavirus strains JNM-6682 and JNM-7354 were isolated from individual yellow pygmy rice rat captured during ecological studies on JUNV in central Argentina from 1987 to 1989 (Mills *et al*., 1991). These isolates were determined to be strains of JUNV based on the results of serological assays. *Machupo virus* strain 9301012 was isolated from a large vesper mouse captured in San Ramon in the Beni Department of Bolivia after a fatal case occurred in that area (WHO, 1995). See Table 4.1 for the histories of the 11 arenaviruses characterized in this chapter.

Virus assay

Spleen or lung tissue homogenates from individual short-tailed cane mouse or Alston's cotton rat were tested for infectious arenavirus as described previously (Fulhorst *et al*., 1999; Weaver *et al*., 2000, 2001). *Junin virus* strains JNM-6682 and JNM-7354 and MACV strain 9301012 were from the virus collection of the Special Pathogens Branch at CDC, Atlanta, Georgia. Propagation of infectious GTOV, JUNV, and MACV was performed by Ms. Mary Louise Milazzo, Dr. Pierre E. Rollin, and Dr. Charles F. Fulhorst inside the BSL-4 laboratory at CDC, Atlanta, Georgia. Similarly, propagation of infectious PIRV was done inside the Keiller Building BSL-3 laboratory. Briefly, monolayer cultures of Vero E6 cells in 25-cm^2 plastic tissue culture flasks were inoculated with diluted virus suspension as previously described (Fulhorst *et al*., 2001). At 7 to 10 days post-inoculation, the virus-infected Vero E6 cell monolayers were lysed in TRIzol® Reagent (Invitrogen Life Technologies, Inc.) or TRI® Reagent BD (Sigma-

Aldrich, MO). The Vero E6 cell lysates were stored at -80°C and then shipped on dry ice to UTMB.

Table 4.1. Histories of 11 arenaviruses isolated from Sigmodontine rodents

		Isolated from				GenBank Accession no.
Species[a]	Strain	Host[b]	State/Province[c]	Country	Year[d]	S segment
GTOV	VHF-1608	Zbre	Apure	Venezuela	1995	AY572560
GTOV	VHF-1750	Zbre	Guarico	Venezuela	1995	AY572559
GTOV	VHF-3990	Zbre	Cojedes	Venezuela	1997	AY576604
GTOV	AV 97021119	Zbre	Portuguesa	Venezuela	1997	AY573922
PIRV	1645	Sals	Apure	Venezuela	1995	AY573921
PIRV	1743	Sals	Guarico	Venezuela	1995	AY575850
PIRV	3945	Sals	Cojedes	Venezuela	1997	AY574571
PIRV	AV 97021016	Sals	Portuguesa	Venezuela	1997	AY573923
JUNV	JNM-6682	Ofla	Santa Fe	Argentina	1988	DQ531486
JUNV	JNM-7354	Ofla	Santa Fe	Argentina	1988	DQ531488
MACV	9301012	Ccal	El Beni	Bolivia	1993	AY924205

[a] JUNV, *Junin virus* (based on serology); MACV, *Machupo virus*; GTOV, *Guanarito virus*; SABV, *Sabiá virus*.
[b] Zbre, *Zygodontomys brevicauda* (short-tailed cane mouse); Sals, *Sigmodon alstoni* (Alston's cotton rat); Ofla, *Oligoryzomys flavescens* (yellow pygmy rice rat); Ccal, *Calomys callosus* (large vesper mouse).
[c]State in Venezuela; province in Argentina; department in Bolivia. Each state/province/department consists of smaller units including towns, municipalities, and localities. Each of the 2 strains of JUNV was isolated from 2 different localities in the province of Santa Fe in central Argentina.
[d]Year in which the specimen was collected.

Preparation of RNA and synthesis of first-strand cDNA

Total RNA was extracted from cell lysates of infected Vero E6 cells following manufacturer's protocols. Reverse transcription on the total RNA was carried out as described previously (Günther *et al*., 2000; Cajimat and Fulhorst, 2004), by using SuperScript II RNase H^- Reverse Transcriptase (Invitrogen Life Technologies, Inc.) in conjunction with oligonucleotide 19C-cons.

PCR amplification and sequencing

The nucleotide sequence of the S genomic segments of GTOV strains VHF-1608, VHF-1750, VHF-3990, and AV 97021119, PIRV strains 1645, 1743, 3945, and AV

97021016, JUNV strains JNM-6682 and JNM-7354, and MACV strain 9301012 were determined from 3 overlapping fragments designated S1, S2, and S3, similar to that described for the human isolates in Chapter 3. Each authentic S genomic segment sequence includes a fragment of the 5' NCR, the complete GPC gene, intergenic region, the complete N protein gene, and a fragment of the 3' NCR. The amplicons were generated from the S segment first-strand cDNA by using Master Taq Kit (Eppendorf North America, Inc.) in conjunction with virus species-specific oligonucleotide primers.

The GTOV S1 fragment was amplified from the first-strand cDNA by using oligonucleotide primer 19C-cons *and* GPC/IR oligonucleotide primers AVGPC16, AVGPC35, or AVNP37; the S2 fragment was amplified by using oligonucleotides AVGPC26 *and* AVNP52 or AVNP53; the S3 fragment was amplified by using oligonucleotides AVNP42 *and* 19C-cons. See "Materials and Methods" section of Chapter 3 for nucleotide sequences of oligonucleotides used in the amplification of GTOV, JUNV, and MACV S genomic segments. The PIRV S1 fragment was amplified from the first-strand cDNA by using oligonucleotides 19C-cons *and* AVGPC57 (5'–AGTCTGTGAGGTTTAGGGCAGCC–3') or AVGPC58 (5'-GTTAATATCCACATG-AGCACAGT-3'); the S2 fragment was amplified by using strain-specific oligonucleotides AVGPC27 (5'- CCTGATGGTGAAGGTTGCCC-3'), AVGPC49 (5'-CAATAACAGGCAGACACAGCCTGC-3'), AVGPC50 (5'-TCTCAGAAACTCTAT-AATGACATGC-3'), or AVGPC59 (5'- GGTCATTGGTCTTCTTCACG-3') in conjunction with AVNP40 (5'-GCCTCGATGTCACACCCC-3'), AVNP54 (5'-AAGAAGAAGGGATCTCCACC-3'), AVNP69 (5'-CAGAGCAAGCGAGACAAT-TCG-3'), or AVNP70 (5'-CACAAAGGTCTTGTGTTAGC-3'); the S3 fragment was amplified by using oligonucleotides AVNP44 (5'-CCCAATTGGGGTGTGACGTCG-AGGC-3') *and* 19C-cons. The JUNV S1 fragment was amplified from the first-strand cDNA by using 19C-cons *and* AVGPC16 or AVGPC97; the S2 fragment was amplified by using oligonucleotides AVGPC73 *and* AVNP99; the S3 fragment was amplified by using oligonucleotide AVNP42 or AVNP87 *and* 19C-cons. Similarly, the MACV S1 fragment was amplified from the first-strand cDNA by using oligonucleotide 19C-cons

and AVGPC16 or AVGPC36; the S2 fragment was amplified by using oligonucleotides AVGPC52 *and* AVNP71; the S3 fragment was amplified by using oligonucleotides AVNP42 *and* 19C-cons.

DNA products of the expected size were purified from agarose gel slices by using the QIAquick Gel Extraction Kit (Qiagen, Inc.) or QIAEXII gel extraction kit (Qiagen, Inc.) and both strands of each amplicon was sequenced directly by using the dye termination cycle sequencing technique (Applied Biosystems, Inc.). The complete S genomic segment sequences were submitted to the GenBank sequence database (see Table 4.1).

Data analysis

Seven strains of LASV isolated from cases of Lassa fever with S genomic segment sequence data were included in the sequence analyses to improve the assessment of "within-species" genetic diversity. The 7 strains of LASV included Josiah from Sierra Leone (Auperin and McCormick, 1989), GA391 from Nigeria (Clegg *et al*., 1991), LP and 803213 from different regions of Nigeria (Bowen *et al*., 2000), AV from the countries of Ivory Coast, Burkina Faso, or Ghana (Günther *et al*., 2000), CSF from Nigeria (Günther *et al*., 2001), and NL from Sierra Leone (Schmitz *et al*., 2002; Vieth *et al*., 2004). Five other strains of LASV with S genomic segment sequence data in the GenBank sequence database were not included because the histories of these viruses were not known: Macenta (GenBank Accession no. AY628201), Z148 (AY628205), Weller (AY628206), Pinneo (AY628207), Acar 3080 (AY628208). Furthermore, in preliminary sequence analyses including the 5 "unknown" strains of LASV, the sequences of the GPC and N protein of strain Pinneo were identical to the homologous sequences of strain LP, and the sequence of the N protein of strain Acar 3080 was identical to the homologous sequences of strains LP and Pinneo. It is possible that these 2 strains were derived from strain LP or that cross contamination occurred in the preparation of the RNAs used in determination of the nucleotide sequences of the S genomic segments.

Construction of multiple amino acid sequence alignments

The predicted amino acid sequences of the GPC genes and the N protein genes of the 11 viruses characterized in this chapter (GTOV strains VHF-1608, VHF-1750, VHF-3990, and AV 97021119; PIRV strains 1645, 1743, 3945, and AV 97021016; JUNV strains JNM-6682 and JNM-7354; MACV strain 9301012) were compared with the homologous sequences of 59 arenaviruses: GTOV strain INH-95551 (GenBank Accession no. NC_005077) and 7 strains from VHF cases described in Chapter 3; PIRV strain VAV-488 (NC_005894); JUNV strains XJ13 (NC_005081), Romero (AY619641), MC2 (D10072), and 10 strains from AHF cases described in Chapter 3; MACV strains Carvallo (AY129248), Mallele (AY619645), 9530537 (AY571959), and 6 strains from BHF cases described in Chapter 3; ALLV strain CLHP-2472 (AY012687); AMAV strain BeAn 70563 (AF512834); CPXV strain BeAn 119303 (AF512832); FLEV strain BeAn 293022 (AF512831); LATV strain MARU10924 (AF485259); OLVV strain 3229-1 (U34248); PARV strain 12056 (AF485261); PICV strains CoAn 3739 (NC_006447) and Munchique (AF081552); SABV strain SPH114202 (NC_006317); TCRV strain TRVL II573 (NC_004293); BCNV strain AV A0070039 (AY924390); TAMV strain W-10777 (AF512828); WWAV strain AV9310135 (AF228063); LCMV strains Armstrong (NC_004294) and WE (M22138); IPPYV strain Dak An B 188d (DQ328877); LASV strains Josiah (NC_004296), GA391 (X52400), LP (AF181853), 803213 (AF181854), AV (AF246121), CSF (AF333969), and NL (AY179173); MOBV strain Acar 3080 (AY342390); MOPV strains 20410 (AY772170), 21366 (M33879), and Mozambique (DQ38874). Collectively, the GPC and N protein sequences of 70 arenaviruses represented the 22 species in the *Arenaviridae*. The amino acid sequences of the GPC and N protein of these viruses were aligned independently using CLUSTAL W1.7 (Thompson *et al*., 1994). The *p*-model distances were calculated from the multiple sequence alignments by using programs in the computer software package MEGA version 3.1 (Kumar *et al*., 2004). Neighbor-joining analyses were carried out on uncorrected *p*-model distances determined from the multiple amino acid sequence alignments.

RESULTS

The multiple amino acid sequence alignments included 12 strains of GTOV, 5 strains of PIRV, 15 strains of JUNV, 10 strains of MACV, 2 strains of PICV, 1 strain each of the 12 other New World arenavirus species, 7 strains of LASV, 3 strains of MOPV, 2 strains of LCMV, and 1 strain each of 2 other Old World species. The GPC and N protein amino acid sequence alignments were 541 and 578 characters in length, respectively. The results of pairwise comparisons of GPC sequences and N protein sequences between strains of a species are summarized in Table 4.2. The results of pairwise comparisons of GPC sequences and N protein sequences among strains of different arenavirus species are summarized in Table 4.3.

Within-species genetic diversity in the *Arenaviridae*

The "within-species" genetic diversity was determined for the arenavirus species that was represented by 5 or more strains. In pairwise comparisons of GPC sequences, the maximum or highest amino acid sequence nonidentity observed within a species was 9.4%, between PIRV strains 1645 and 1743 (Table 4.2). Similarly, in pairwise comparisons of N protein sequences, the maximum "within-species" sequence nonidentity was 11.6%, between LASV strains CSF and LP (Table 4.2).

Table 4.2. Within-species genetic diversity in the *Arenaviridae*

	Amino acid sequence nonidentity (%)[b]			
Species[a]	GPC (range)	GPC (mean)	N protein (range)	N protein (mean)
GTOV	0.0 – 8.8	4.3	0.2 – 4.5	2.4
PIRV	4.7 – 9.4	7.1	4.5 – 9.3	7.0
JUNV	0.0 – 2.5	1.0	0.0 – 3.9	1.3
MACV	0.0 – 6.0	3.6	0.0 – 2.7	1.6
LASV	2.4 – 8.6	6.9	3.3 – 11.6	9.0

[a]GTOV, *Guanarito virus* (12 strains); PIRV, *Pirital virus* (5 strains); JUNV, *Junin virus* (15 strains); MACV, *Machupo virus* (10 strains); LASV, *Lassa virus* (7 strains).
[b]GPC, glycoprotein precursor open reading frame; N protein, nucleocapsid protein open reading frame.

Between-species genetic diversity in the *Arenaviridae*

In pairwise comparisons of 70 GPC sequences, the minimum or lowest amino acid sequence nonidentity observed between strains of different arenaviral species was 15.8%, between ALLV strain CLHP-2472 and FLEV strain BeAn 293022 (Table 4.3). Similarly, in pairwise comparisons of N protein sequences, the lowest "between-species" diversity 11.2%, between JUNV strain XJ13 and MACV strain 9430084. The highest sequence nonidentity between species in both GPC and N protein was observed between a New World species and an Old World species. Therefore, arenaviruses that differ by more than 15.8% in the GPC and more than 11.2% in the N protein could belong to different species. Collectively, the results of pairwise comparisons of GPC and N protein sequences indicate that a novel arenavirus species must exhibit at least 11% sequence nonidentity in the GPC and in the N protein in comparison with strains of known arenavirus species.

Phylogenetic relationships of the Tacaribe serocomplex viruses

The predicted amino acid sequences of the GPC gene and N protein gene of the 11 arenaviruses characterized in this chapter were aligned with the homologous sequences of 44 other Tacaribe serocomplex viruses, LASV strain Josiah, and LCMV strain WE. The 6 other strains of LASV and 6 strains of the other Old World species were not included in the phylogenetic analyses; the phylogeny of the Old World arenaviruses has recently been described elsewhere (Emonet *et al*., 2006).

Table 4.3. Between-species genetic diversity in the *Arenaviridae*

Species[a]	Amino acid sequence nonidentity (%)[b]			
	GPC (min)	GPC (max)	N protein (min)	N protein (max)
lymphocytic choriomeningitis-Lassa serocomplex				
Lymphocytic choriomeningitis virus	40.0	66.2	35.4	51.1
Ippy virus	26.7	63.0	29.6	54.3
Lassa virus	21.4	62.8	24.3	54.2
Mobala virus	20.4	61.7	21.5	53.5
Mopeia virus	20.4	61.8	21.5	51.6
Tacaribe serocomplex				
Allpahuayo virus	**15.8**	61.7	23.4	50.2
Flexal virus	**15.8**	60.9	21.9	52.3
Parana virus	17.9	62.1	21.9	50.3
Pichindé virus	17.4	62.2	23.4	51.2
Pirital virus	17.4	61.8	29.5	50.9
Amapari virus	29.2	63.5	14.6	50.4
Cupixi virus	30.3	62.6	16.4	49.9
Guanarito virus	29.2	63.8	14.6	51.1
Junin virus	28.2	64.7	**11.2**	53.3
Machupo virus	28.2	66.2	**11.2**	51.6
Sabiá virus	44.3	66.0	27.8	50.8
Tacaribe virus	32.8	65.4	19.9	51.1
Latino virus	20.4	59.8	19.0	50.8
Oliveros virus	20.4	58.7	19.0	51.0
Bear Canyon virus	34.1	62.4	18.3	51.5
Tamiami virus	34.1	64.5	19.9	54.3
Whitewater Arroyo virus	38.5	65.0	18.3	51.2

[a]A total of 70 arenaviruses with complete glycoprotein precursor (GPC) and nucleocapsid (N) protein gene sequences were compared. Arenaviral species with 5 or more strains included *Lassa virus*, *Guanarito virus*, *Junin virus*, *Machupo virus*, and *Pirital virus*.
[b]Between-species sequence nonidentities represent *p*-model distances calculated from the multiple amino acid sequence alignments. Min (minimum) and max (maximum) correspond to the lowest and highest amino acid sequence nonidentity exhibited between strains of different arenaviral species, respectively. The GPC and N protein amino acid sequence alignments were 541 and 578 characters in length, respectively.

The neighbor-joining (NJ) analysis of *p*-model distances generated from the GPC sequence alignment grouped the 4 strains of GTOV isolated from short-tailed cane mice with the 8 strains isolated from human cases (Figure 4.1A). In the GTOV lineage containing 12 strains, the human strains except S-56764 grouped together in a minor

lineage that was well-supported (97% bootstrap support). The rodent strain AV 97021119 grouped together with the human strain S-56764, both of which were isolated from different localities within the state of Portuguesa. The relationships of the 3 rodent strains (VHF1608, VHF-1750, VHF-3990) was not resolved. Similarly, the 5 strains of PIRV were monophyletic with a bootstrap support of 100%. The relationships of PIRV and the 4 South American lineage A species (PICV, ALLV, FLEV, and PARV) were not resolved. Within the PIRV lineage, strain 1645 from the state of Apure was deeply rooted with respect to the 4 other strains. In the JUNV lineage, the 2 strains isolated from from *O. flavescens*, JNM-6682 and JNM-7354, grouped with human strains of JUNV. Altogether, the 15 strains of JUNV were monophyletic although most of the relationships between strains could not be resolved. Strain MC2 isolated from a species of *Calomys*, was deeply rooted to the other 14 JUNV strains. Consistent with GTOV, PIRV, and JUNV, all 10 strains of MACV were monophyletic; the MACV lineage was sister to the JUNV lineage. Within the MACV lineage, the lone isolate from rodent (strain 9301012) grouped with strain Chicava, both of which were isolated in the same year and in the same locality. Strain 9530537 was deeply rooted to all other 9 MACV strains. This strain was isolated in 1995 from the Department of La Paz, outside the known BHF-endemic region of Bolivia.

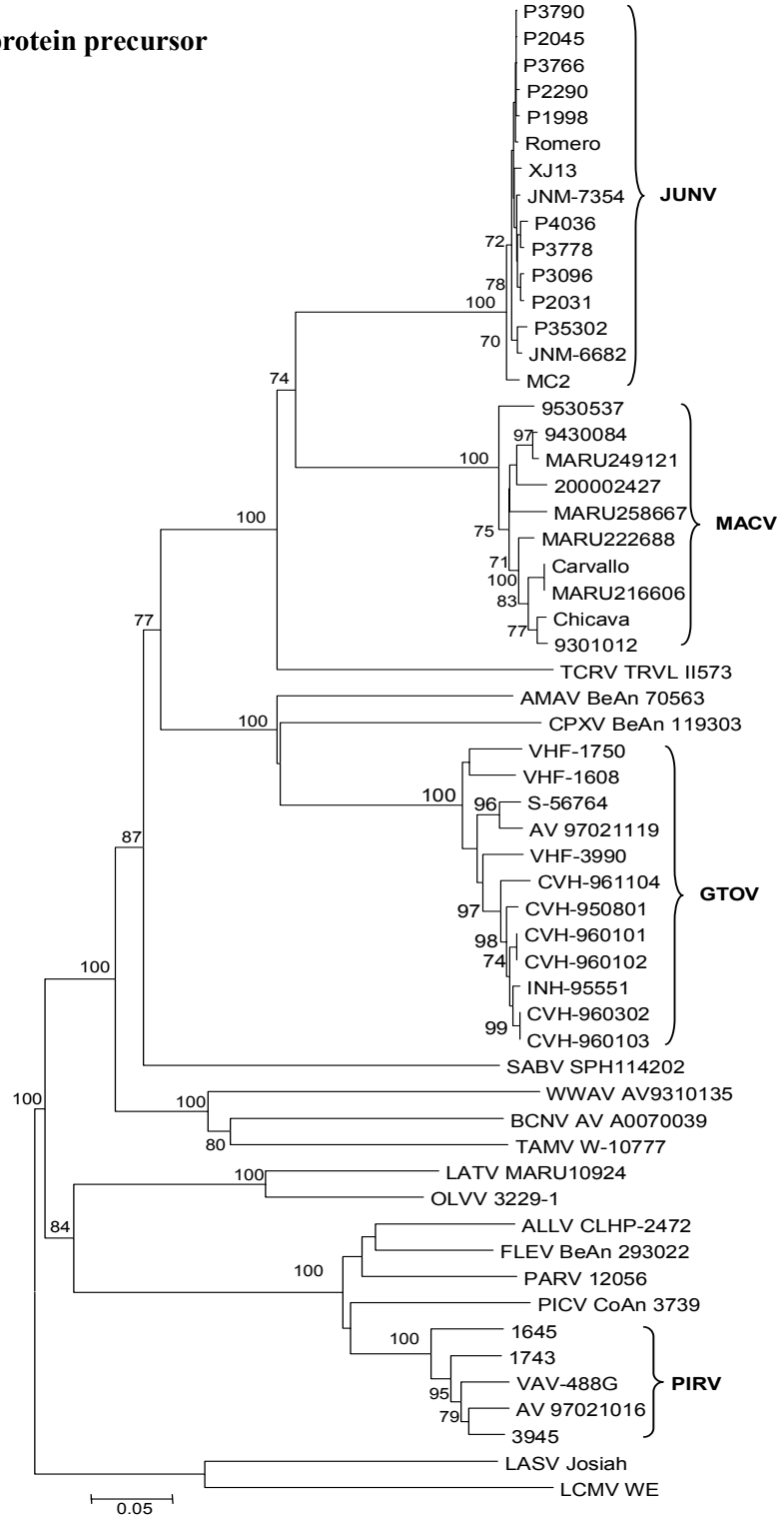
The NJ analysis of *p*-model distances generated from the N protein sequence alignment grouped the 4 GTOV strains isolated from rodents with the 8 strains isolated from VHF cases (Figure 4.1B). All 12 strains of GTOV were monophyletic in a lineage that was sister to AMAV. All human strains except S-56764 grouped in a minor lineage supported by 80% bootstrap support. Similar to the GPC sequence analysis, the rodent strain AV 97021119 grouped together with strain S-56764. Strains VHF-1608 and VHF-1750 comprised a minor lineage separate from the other 10 GTOV strains. The 5 strains of PIRV were monophyletic in N protein sequence analysis and the PIRV lineage was separate from the lineage containing the 4 other South American lineage A species. Within the PIRV lineage, strain 1645 was deeply rooted to the other 4 strains. The 2 isolates from *O. flavescens* grouped with human strains of JUNV. Altogether, the 15

strains of JUNV were monophyletic with strain MC2 being deeply rooted to the other 14 strains. The rest of the relationships within the JUNV lineage were not resolved. In the same N protein sequence analysis, all 10 strains of MACV were monophyletic, and the MACV lineage was sister to the lineage containing the 15 strains of JUNV. Except for the monophyly of the lone rodent strain and the Chicava strain and the monophyly of strains Carvallo and MARU216606, the rest of the relationships within the MACV lineage were not resolved. Strain 9530537 again was deeply rooted to all other 9 strains.

DISCUSSION

Several properties typically are used to distinguish viral species within a genus or virus family. The properties or characters used usually are not identical between genera within a virus family and between virus families. The ICTV study group for each virus family selects the criteria to be used for species demarcation. Genetic sequence differences have now become an important element of these species demarcation criteria as evident in the taxonomy of many virus families. For example, the genetic criteria for species demarcation in the genus *Hantavirus* (family *Bunyaviridae*) requires 7% (or greater) amino acid sequence nonidentity in complete GPC and complete N protein sequences to distinguish hantaviral species. The use of gene sequence differences as part of species demarcation criteria is particularly useful in the context of classifying highly pathogenic viruses such as the South American hemorrhagic fever arenaviruses; determining the antigenic relationships (i.e., serotype) of these viruses by neutralization tests can only be done in a high containment facility, requires immune reagents that are not available, cumbersome and requires a considerable length of time to complete. Genetic criteria also is useful in the context of a species complex, that is – members of a species that are not distinguishable by physical properties or other phenotypic characteristics and can only be separated on the basis of differences in genome sequences.

A. Glycoprotein precursor



B. Nucleocapsid protein

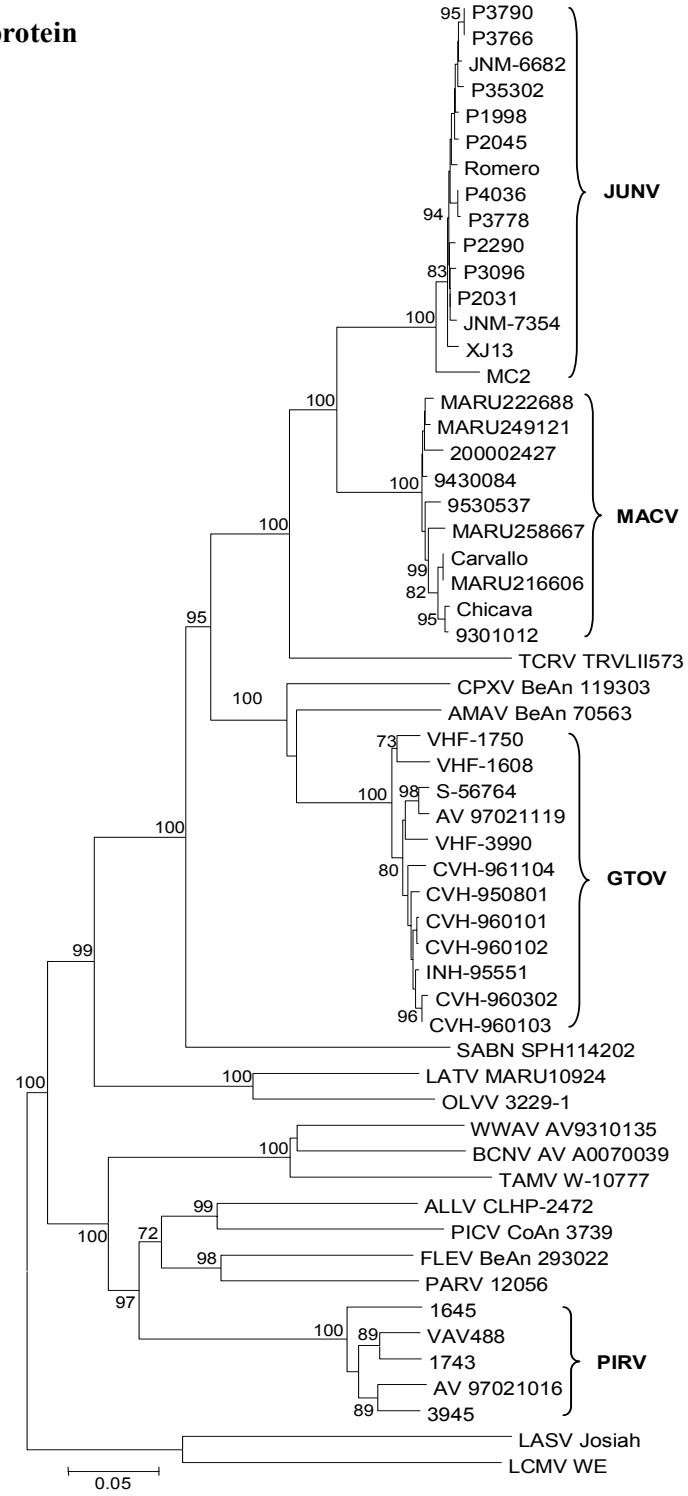


Figure 4.1. Phylogenetic relationships of the Tacaribe serocomplex viruses based on neighbor-joining analysis of (A) glycoprotein precursor sequences and (B) nucleocapsid protein sequences. The scale bar indicates a sequence divergence of 0.05. The numerical value at the node indicates the percentage of 1000 bootstrap replicates that supported the interior branch. Bootstrap support values less than 70% are not listed. The branch labels include viral species and strain; multiple strains are bracketed with the viral species label. ALLV, *Allpahuayo virus*; AMAV, *Amapari virus*; BCNV, *Bear Canyon virus*; CPXV, *Cupixi virus*; FLEV, *Flexal virus*; GTOV, *Guanarito virus*; JUNV, *Junin virus*; LATV, *Latino virus*; LCMV, *lymphocytic choriomeningitis virus*; MACV, *Machupo virus*; OLVV, *Oliveros virus*; PARV, *Parana virus*; PICV, *Pichindé virus*; PIRV, *Pirital virus*; SABV, *Sabiá virus*; TCRV, *Tacaribe virus*; TAMV, *Tamiami virus*; WWAV, *Whitewater Arroyo virus*. The LASV strain Josiah and LCMV strain WE are Old World arenaviruses and were included in the analyses to infer the ancestral node within the group of New World arenaviruses.

The criteria used by ICTV for species demarcation in the *Arenaviridae* is highly similar to that used in the genus *Hantavirus*. The species demarcation criteria for the genus *Hantavirus* are 1) the virus must occupy a unique ecological niche (i.e., primary rodent reservoir species), 2) exhibit at least 7% difference in GPC and N protein amino acid sequences, 3) the virus must represent a unique serotype (i.e., four-fold difference in two-way cross neutralization tests), and 4) the virus must not form reassortants with another species in nature (Nichol *et al*., 2005). Similar to the arenaviruses, the natural reservoirs of the hantaviruses are members of the rodent families Muridae and Cricetidae. A number of hantaviruses are etiologic agents of severe disease in humans; for example, *Hantaan virus* in Asia is a causative agent of hemorrhagic fever with renal syndrome, and *Sin Nombre virus* in North America is a causative agent of hantavirus pulmonary syndrome (Nichol, 2001). Historically, the results of serological assays were used to establish the taxonomy of the genus *Hantavirus*. Because of these similarities, we modeled our genetic criteria for species demarcation in the *Arenaviridae* based on the criteria set forth by the ICTV for the genus *Hantavirus*.

For the genus *Arenavirus*, the genetic criteria is stated as "significant amino acid sequence differences" (Salvato *et al*., 2005); it does not specify a gene and there is no numerical measurement as to what degree of sequence differences is considered significant. In the study described in this chapter, we utilized sequence data previously generated (in the current and previous chapters) as well as sequence data available from

the GenBank sequence database to assess "within-species" genetic diversity and "between-species" genetic diversity in the *Arenaviridae* to come up with a lower limit or cutoff value for the genetic criteria. Knowledge of the extent of genetic diversity within the *Arenaviridae* will enable us, the ICTV study group, and viral taxonomists to provide a consistent guideline to be used in the classification of arenaviruses.

The current genetic criteria for species demarcation in the genus *Hantavirus* require at least 7.0% amino acid sequence difference both in the GPC and in the N protein between hantaviral species. The results of the present study indicate that a 7.0% amino acid sequence nonidentity in the GPC and in the N protein as lower limit is too low to distinguish between arenaviral species. The results of pairwise comparisons of GPC sequences and N protein sequences among 70 strains of 22 arenaviral species indicate that a lower limit of 11.2% in both the GPC and in the N protein must be exhibited to distinguish between species. This value was the minimum sequence nonidentity obtained between a strain of JUNV and a strain of MACV, the 2 most closely related arenaviruses on the basis of serological assays and genetic analyses. The highest sequence nonidentity in the GPC among strains of a species was 9.4% and was exhibited between strains of PIRV. The highest sequence nonidentity in the N protein within a species was 11.6%, between strains of LASV. Among the New World arenaviral species, within-species diversity in the N protein was highest between strains of PIRV. Thus, a lower limit of 11.2% sequence nonidentity in both GPC and N protein sequences is sufficient to define species boundaries among the Tacaribe serocomplex viruses.

In contrast, this lower limit of 11.2% sequence nonidentity may not be sufficient to distinguish between members of the lymphocytic choriomeningitis-Lassa serocomplex. The highest sequence nonidentity observed between strains of LASV in the N protein (11.6%) exceeded the minimum sequence nonidentity between JUNV and MACV (11.2%). Viruses known as LASV in association with different *Mastomys* species are widely distributed in 4 or more countries of west Africa (McCormick *et al*., 1987). These are serologically and genetically distinguishable from the other Old World species

indigenous to Africa. The extent of sequence diversity among viruses known as LASV suggests that multiple species may exist within what is known as LASV (Bowen *et al*., 2000). However, the lack of data on the rodent host associations of LASV strains with available sequence data together with the unresolved phylogeography of the natural rodent host makes it difficult to evaluate this hypothesis.

The results of a previous study (Fulhorst *et al*., 2001) indicated that arenaviruses genetically most closely related to WWAV are widely geographically distributed throughout the southwestern United States. Pairwise comparisons of the deduced amino acid sequences of a 518-nt fragment of the N protein gene of 12 strains and WWAV strain AV9310135 revealed as much as 15.1% sequence nonidentity (*p*-model distances) between these viruses. Even with this degree of genetic variability, they are genetically distinguishable from BCNV and TAMV. All these strains were isolated from 4 different species of woodrats (genus *Neotoma*). The extent of genetic variability among these viruses known as WWAV, together with the co-existence of their rodent hosts in nature suggests that they may be part of a WWAV species complex similar to LASV in west Africa. We propose to use the genetic criteria in conjunction with the 3 other criteria (specific host association, defined geographic range, human pathogen or not) and if available, serological data to define these viruses naturally associated with *Neotoma* species.

Chapter 5: *Catarina virus*, an arenaviral species principally associated with *Neotoma micropus* (southern plains woodrat) in Texas

INTRODUCTION

The virus family *Arenaviridae*, genus *Arenavirus* comprises 2 serocomplexes (Salvato *et al*., 2005). The lymphocytic choriomeningitis-Lassa (Old World) complex includes *Lassa virus* (LASV), *lymphocytic choriomeningitis virus* (LCMV), *Ippy virus* (IPPV), *Mobala virus* (MOBV), and *Mopeia virus* (MOPV). The Tacaribe (New World) complex includes *Bear Canyon virus* (BCNV), *Tamiami virus* (TAMV), and *Whitewater Arroyo virus* (WWAV) in North America, *Tacaribe virus* (TCRV) on Trinidad in the Caribbean Sea, and *Allpahuayo virus* (ALLV), *Amapari virus* (AMAV), *Cupixi virus* (CPXV), *Flexal virus* (FLEV), *Guanarito virus* (GTOV), *Junin virus* (JUNV), *Latino virus* (LATV), *Machupo virus* (MACV), *Oliveros virus* (OLVV), *Parana virus* (PARV), *Pichindé virus* (PICV), *Pirital virus* (PIRV), and *Sabiá virus* (SABV) in South America.

Specific members of the subfamilies Neotominae and Sigmodontinae in the rodent family Cricetidae (Musser and Carleton, 2005) are the principal hosts of the New World arenaviruses for which natural host relationships have been well characterized. For example, the white-throated woodrat (*Neotoma albigula*) in northwestern New Mexico is the principal host of WWAV (Fulhorst *et al*., 1996), the hispid cotton rat (*Sigmodon hispidus*) in southern Florida is the principal host of TAMV (Calisher *et al*., 1970; Jennings *et al*., 1970), Alston's cotton rat (*Sigmodon alstoni*) in western Venezuela is the principal host of PIRV (Fulhorst *et al*., 1997, 1999), the drylands vesper mouse (*Calomys musculinus*) in northern Argentina is the principal host of JUNV (Mills *et al*., 1992), and the large vesper mouse (*Calomys callosus*) in northeastern Bolivia is the principal host of MACV (Johnson *et al*., 1966). The white-throated woodrat is a neotomine rodent whereas the hispid cotton rat, Alston's cotton rat, drylands vesper mouse, and large vesper mouse are sigmodontine rodents (Musser and Carleton, 2005).

A previous study established that the southern plains woodrat (*Neotoma micropus*) in southern Texas is the principal host of an arenavirus that is phylogenetically more closely related to WWAV than to TAMV, TCRV, or the South American arenaviral species (Fulhorst *et al*., 2002b). The purpose of the present study was to define better the taxonomical and phylogenetical relationship of the arenavirus associated with the southern plains woodrat in southern Texas to WWAV and the other New World arenaviruses.

MATERIALS AND METHODS

Viruses

Strains AV A0400135 and AV A0400212 originally were isolated from southern plains woodrats captured in July 1999 at a site located approximately 18.7 km east of the town of Catarina in Dimmit County, Texas (Fulhorst *et al*., 2002b). The results of a comparison of predicted amino acid sequences in a previous study (Fulhorst *et al*., 2002b) indicated that the primary structure of the N protein of AV A0400135 is different from the primary structure of the N protein of AV A0400212.

PCR amplification and sequencing the S genomic segments of strains AV A0400135 and AV A0400212

The nucleotide sequence of a 3288-nt fragment of the S genomic segment of AV A0400135 and the nucleotide sequence of a 3287-nt fragment of the S genomic segment of AV A0400212 were determined in a manner similar to that described previously (Charrel *et al*., 2001). Each sequence includes a fragment of the 5' NCR, the complete GPC gene (1452-nt), intergenic region (71- or 72-nt), the complete N protein gene (1686-nt), and a fragment of the 3' NCR. Total RNA was isolated from monolayers of infected Vero E6 cells, using TRIzol® Reagent (Invitrogen Life Technologies, Inc.). First-strand cDNA was generated from the arenaviral RNA by using SuperScript II RNase H^- Reverse Transcriptase (Invitrogen Life Technologies, Inc.) in conjunction with oligonucleotide

19C-cons. This oligonucleotide is a derivative of oligonucleotide ARE3'-END (Gonzalez *et al*., 1995) and was expected to anneal to the 19-nt fragment at the extreme 3' end of the S segment and the 19-nt fragment at the extreme 3' end of the replicative intermediate synthesized from the S segment (Auperin *et al*., 1982a). Amplicons (PCR products) were generated from three overlapping fragments of the first-strand cDNA by using the Master Taq Kit (Eppendorf) in conjunction with 19C-cons *and* AVGPC14 (5'-GGACAGCCY-TCRCCRATKATGTGTCTGTG-3'), AVGPC54 (5'-ATCTCATCTCTGAAATCCTGA-G-3') *and* AVNP72 (5'-GTTGATGTGAAGCTAAGTGC-3'), and AVNP13 (5'-GTTGTKTCWGGYTCYCTGAA-3') *and* 19C-cons. Oligonucleotides 19C-cons and AVGPC14 flank a 1426-nt fragment of the S segment that extends from within the 5' NCR; AVGPC54 and AVNP72 flank a 598-nt fragment (strain AV A0400135) or 597-nt fragment (strain AV A0400212) that extends from within the GPC gene, across the intergenic region, and through the stop codon of the N protein gene; AVNP13 and 19C-cons flank a 1726-nt fragment that extends from within the N protein gene, through the start codon of the N protein gene, and into the 3' NCR. Amplicons of the expected size were purified from agarose gel slices by using the QIAquick Gel Extraction Kit (Qiagen Inc.). Both strands of each purified amplicon were sequenced directly, using the dye termination cycle sequencing technique (Applied Biosystems Inc.). The sequence of the 3288-nt fragment of the S segment of AV A0400135 and the sequence of the 3287-nt fragment of the S segment of AV A0400212 were deposited into the GenBank nucleotide sequence database under Accession nos. DQ865244 and DQ865245, respectively.

Sequence analyses

The amino acid sequences of the GPC and N proteins of AV A0400135 and AV A0400212 were compared to the homologous sequences of BCNV strain AV A0060209, TAMV strain W-10777, WWAV strain AV 9310135, ALLV strain CLHP-2472, AMAV strain BeAn 70563, CPXV strain BeAn 119303, FLEV strain BeAn 293022, GTOV strain INH-95551, JUNV strains XJ13, MC2, and Romero, LATV strain MARU 10924, MACV strains Carvallo, Chicava, Mallele, and 9530537, OLVV strain 3229-1, PARV

strain 12056, PICV strain An 3739, PIRV strain VAV-488, SABV strain SPH 114202, TCRV strain TRVL II573, and LCMV strain WE (GenBank Accession Nos. AF512833, AF512828, AF228063, AY012687, AF512834, AF512832, AF512831, NC_005077, NC_005081, D10072, AY619641, AF485259, NC_005078, AY624355, AY619645, AY571959, U34248, AF485261, NC_006447, NC_005894, NC_006317, NC_004293, and M22138, respectively). The LCMV strain WE was included in the phylogenetic analyses to enable inference of the ancestral node within the group of New World arenaviruses. The GPC and N protein amino acid sequences were aligned independently, using the computer program CLUSTAL W1.7 (Thompson *et al*., 1994). The GPC and N protein gene nucleotide sequence alignments were constructed manually based on alignments of the GPC and N protein amino acid sequences, respectively. The analyses of the multiple amino acid sequence alignments were done by using programs in the computer software package MEGA version 2.1 (Kumar *et al*., 2001). The neighbor-joining analyses were carried out on uncorrected p-model distances generated from the multiple amino acid sequence alignments. Bootstrap support for the results of each neighbor-joining analysis was based on 1000 repetitions of the heuristic search, with random resampling of the data (Felsenstein, 1985).

The analyses of the multiple nucleotide sequence alignments were performed in the laboratory of Dr. Robert Bradley at Texas Tech University, Lubbock, Texas and were done by using MrBayes 3.1.2 and other programs in the computer software package PAUP (Huelsenbeck and Ronquist, 2001; Swofford, 2002). A GTR+I+G model with a site-specific gamma distribution and sites partitioned by codon was used with the following options: 4 Markov-chains, 1 million generations, sample frequency = every 1,000th generation. The likelihood scores, convergence statistics, and potential scale reduction factors were reviewed, and then the first 1,000 trees were discarded. A consensus tree (50% majority rule) was constructed from the remaining trees, and clade probability values were generated to assess support for the nodes within the consensus tree.

RESULTS

The GPC amino acid sequence alignment was 560 characters in length and the N protein amino acid sequence alignment was 576 characters in length. Sequence nonidentities (uncorrected *p*-model distances) between the GPC and between the N proteins of AV A0400135 and AV A0400212 were 4.8% and 1.6%, respectively (Table 5.1). Sequence nonidentities between the GPC and between the N proteins of JUNV strains XJ13, MC2, and Romero ranged from 0.8% to 1.6% and from 1.2% to 3.9%, respectively. Similarly, sequence nonidentities between the GPC and between the N proteins of MACV strains Carvallo, Chicava, Mallele and 9530537 ranged from 1.6% to 5.6% and from 0.9% to 2.3%, respectively. Thus, AV A0400135 and AV A0400212 are conspecific, that is -- strains of the same arenaviral species.

The GPC of AV A0400135 and AV A0400212 exhibited the lowest sequence nonidentity (33.1% and 33.5%, respectively) with the GPC of WWAV strain AV 9310135 (Table 5.1). Further, the N proteins of AV A0400135 and AV A0400212 exhibited the lowest sequence nonidentity (13.5% and 13.3%, respectively) with the N protein of WWAV strain AV 9310135 (Table 5.1).

Sequence nonidentities between the GPC of strains of different New World arenaviral species ranged from 15.8% (ALLV strain CLHP-2472 and FLEV strain BeAn 293022) to 60.3% (AMAV strain BeAn 70563 and PIRV strain VAV-488) and sequence nonidentities between the N proteins of strains of different New World arenaviral species ranged from 11.9% (JUNV strain XJ13 and MACV strain 9530537) to 47.5% (AMAV strain BeAn 70563 and TAMV strain W-10777) (Table 5.1). Thus, the arenaviral species represented by A0400135 and AV A0400212 should be considered distinct from WWAV and other New World arenaviral species. The name "*Catarina virus*" is proposed to distinguish the arenaviral species associated with *N. micropus* in southern Texas from all other arenaviral species.

Table 5.1. Amino acid sequence nonidentities in the glycoprotein precursor and nucleocapsid protein of 19 New World arenaviruses[a]

Species[b]	CTNV-135	CTNV-212	BCNV	TAMV	WWAV	ALLV	AMAV	CPXV	FLEV	GTOV	JUNV	LATV	MACV	OLVV	PARV	PICV	PIRV	SABV	TCRV
CTNV-135	--	4.8	35.7	35.7	33.1	54.8	52.7	47.8	55.4	50.2	49.1	53.7	49.4	51.7	55.6	57.3	57.0	46.1	50.4
CTNV-212	1.6	--	35.7	36.1	33.5	55.2	53.1	48.4	56.3	50.6	49.4	54.3	50.0	52.3	56.1	57.7	57.0	47.0	50.9
BCNV	18.1	17.8	--	34.7	40.2	57.7	49.6	45.5	56.2	47.8	48.0	53.4	49.5	53.2	58.7	58.2	57.3	46.6	49.5
TAMV	18.5	18.1	21.7	--	38.3	56.6	50.5	46.9	55.1	48.5	49.4	51.3	49.6	49.9	56.6	57.4	56.3	47.6	51.6
WWAV	13.5	13.3	18.7	19.9	--	57.0	53.6	48.9	57.8	51.2	50.9	55.0	49.4	52.3	58.0	58.6	58.1	48.3	51.7
ALLV	38.4	38.0	35.7	40.2	36.6	--	59.3	59.0	15.8	58.9	58.8	49.9	58.8	51.4	17.9	21.0	18.5	55.4	57.9
AMAV	46.4	46.4	45.9	47.5	45.0	42.2	--	30.4	57.8	30.1	42.9	55.6	43.4	55.4	59.3	58.9	60.3	45.6	46.4
CPXV	46.4	46.4	47.0	47.0	46.1	42.2	17.1	--	56.9	31.3	46.1	51.9	45.9	52.0	58.4	56.5	56.9	43.3	46.2
FLEV	35.8	35.8	34.7	38.1	37.4	24.9	42.7	44.0	--	55.3	56.9	48.3	57.5	50.6	18.5	23.4	23.7	55.4	56.0
GTOV	44.8	44.8	45.3	46.4	44.6	43.1	14.6	17.1	41.3	--	43.6	52.0	45.1	50.8	56.6	59.0	56.8	45.3	46.5
JUNV	46.1	46.1	46.1	46.2	45.2	42.2	26.3	27.1	41.8	25.2	--	55.1	28.9	53.7	59.5	59.7	58.5	45.0	32.2
LATV	42.8	43.1	41.5	43.3	41.7	40.3	36.9	35.4	42.9	35.8	38.8	--	53.8	20.4	51.5	51.4	50.4	50.6	54.5
MACV	45.7	45.9	45.5	47.0	44.6	41.7	25.7	25.9	43.1	23.7	12.2	37.0	--	51.4	58.6	59.3	58.9	44.2	35.0
OLVV	44.0	44.0	43.1	42.9	43.8	41.0	37.3	36.1	41.9	37.9	39.2	19.0	37.8	--	52.0	53.1	52.3	50.9	53.6
PARV	38.5	38.5	34.9	38.1	39.2	28.0	42.9	43.6	21.3	41.8	42.2	41.2	43.4	41.9	--	21.0	21.7	56.8	58.5
PICV	36.6	36.1	35.7	37.0	35.5	23.2	42.4	44.7	29.2	42.7	43.6	42.3	44.3	42.8	32.4	--	20.5	57.7	59.2
PIRV	37.7	37.3	36.6	38.2	37.0	29.5	43.3	44.0	31.7	41.7	42.9	40.1	43.4	40.8	32.1	29.5	--	57.5	57.9
SABV	44.7	44.9	44.7	45.1	44.4	43.2	27.9	29.1	43.2	29.8	30.1	37.9	28.5	38.4	43.2	41.9	43.5	--	47.7
TCRV	46.4	46.2	45.9	47.1	45.7	40.9	29.8	30.7	47.6	28.7	21.3	39.3	20.4	39.6	42.9	43.3	42.7	33.2	--

[a]Numbers above and below the diagonal line are the amino acid sequence nonidentities between the glycoprotein precursor and between the nucleocapsid proteins, respectively.
[b]CTNV-135, *Catarina virus* strain AV A0400135; CTNV-212, *Catarina virus* strain AV A0400212; BCNV, *Bear Canyon virus* strain AV A0060209; TAMV, *Tamiami virus* strain W-10777; WWAV, *Whitewater Arroyo virus* strain AV 9310135; ALLV, *Allpahuayo virus* strain CLHP-2472; AMAV, *Amapari virus* strain BeAn 70563;

CPXV, *Cupixi virus* strain BeAn 119303; FLEV, *Flexal virus* strain BeAn 293022; GTOV, *Guanarito virus* strain INH-95551; JUNV, *Junin virus* strain XJ13; LATV, *Latino virus* strain MARU 10924; MACV, *Machupo virus* strain Carvallo; OLVV, *Oliveros virus* strain 3229-1; PARV, *Parana virus* strain 12056; PICV, *Pichindé virus* strain An 3739; PIRV, *Pirital virus* strain VAV-488; SABV, *Sabiá virus* strain SPH 114202; TCRV, *Tacaribe virus* strain TRVL II573.

The neighbor-joining analysis of uncorrected *p*-model distances generated from the alignment of GPC sequences placed *Catarina virus* (CTNV) in a sister relationship to WWAV and separated the North American species (BCNV, CTNV, TAMV, and WWAV) from TCRV and the South American species (Figure 5.1A). The neighbor-joining analysis of uncorrected *p*-model distances generated from the alignment of N protein sequences also placed CTNV in a sister relationship to WWAV and separated the North American species from TCRV and the South American species (Figure 5.1B). Bootstrap support for monophyly of CTNV and WWAV and for monophyly of the 4 North American arenaviral species in the phylograms generated from the GPC sequence data was greater than 95.0%. Bootstrap support for monophyly of CTNV and WWAV and for monophyly of the 4 North American arenaviral species in the phylograms generated from the N protein sequence data was greater than 97.0%.

The 2 trees generated by simultaneous Bayesian analyses of the GPC gene nucleotide sequence data were identical (data not shown). The 2 trees generated by simultaneous Bayesian analyses of the N protein gene nucleotide sequence data also were identical (data not shown). In the analyses of the GPC gene sequences, CTNV was sister to WWAV, the CTNV-WWAV clade was sister to BCNV, the BCNV-CTNV-WWAV clade was sister to TAMV, and the North American (TAMV-BCNV-CTNV-WWAV) clade was sister to a clade that comprised AMAV, CPXV, GTOV, JUNV, MACV, SABV, and TCRV. The clade probability values for monophyly of CTNV and WWAV and monophyly of the North American viruses was 100; but, the clade probability value for the sister relationship between BCNV and the CTNV-WWAV clade was less than 95. In the analyses of the N protein gene sequences, CTNV was sister to WWAV, the CTNV-WWAV clade was sister to TAMV, the TAMV-CTNV-WWAV clade was sister to BCNV, and the North American (BCNV-TAMV-CTNV-WWAV) clade was sister to a clade that comprised ALLV, FLEV, PARV, PICV, and PIRV. The clade probability values for the North American clade and for each node within the North American clade were 100.

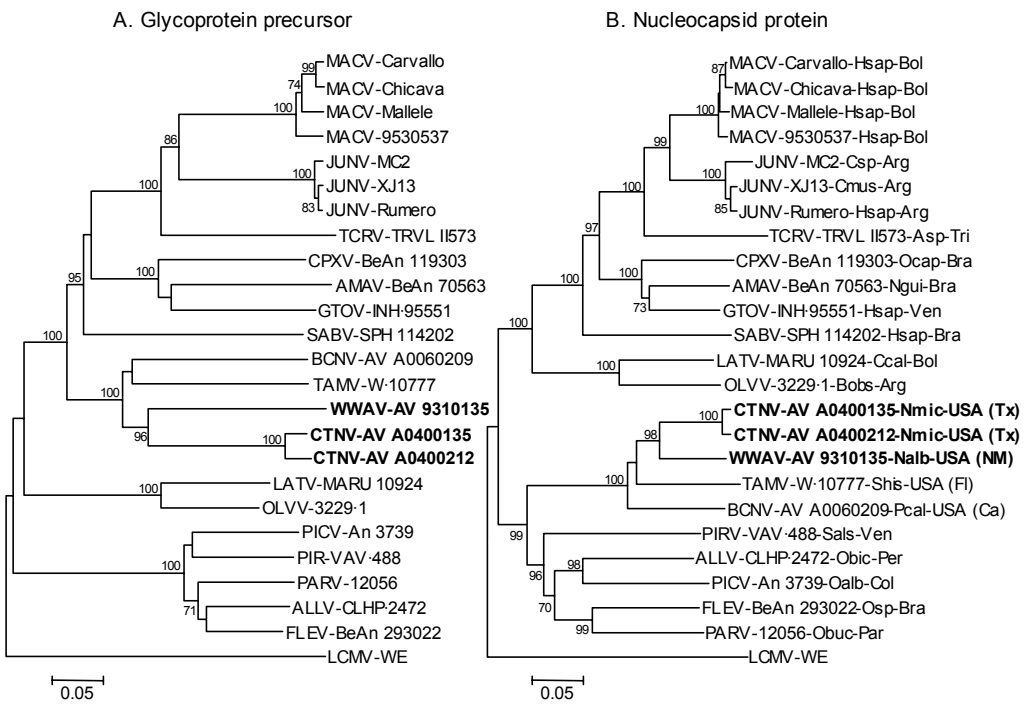


Figure 5.1. Phylogenetic relationships among 24 New World arenaviruses based on neighbor-joining analyses of (A) full-length glycoprotein precursor sequences and (B) full-length nucleocapsid protein sequences. The scale bar indicates a sequence divergence of 0.05. The numerical value at the node indicates the percentage of 1000 bootstrap replicates that supported the interior branch. Bootstrap support values less than 70% are not listed. The branch labels in B include (in the following order) viral species, strain, host species and country. ALLV, *Allpahuayo virus*; AMAV, *Amapari virus*; BCNV, *Bear Canyon virus*; CPXV, *Cupixi virus*; CTNV, *Catarina virus*; FLEV, *Flexal virus*; GTOV, *Guanarito virus*; JUNV, *Junin virus*; LATV, *Latino virus*; LCMV, *lymphocytic choriomeningitis virus*; MACV, *Machupo virus*; OLVV, *Oliveros virus*; PARV, *Parana virus*; PICV, *Pichindé virus*; PIRV, *Pirital virus*; SABV, *Sabiá virus*; TCRV, *Tacaribe virus*; TAMV, *Tamiami virus*; WWAV, *Whitewater Arroyo virus*. Arg, Argentina; Bol, Bolivia; Bra, Brazil; Col, Colombia; Par, Paraguay; Per, Peru; Tri, Trinidad; USA, United States of America (Ca, California; Fl, Florida; NM, New Mexico; Tx, Texas); Ven, Venezuela. Asp, *Artibeus* species (frugivorous bats); Bobs, *Bolomys obscurus* (dark bolo mouse); Ccal, *Calomys callosus* (large vesper mouse); Cmus, *Calomys musculinus* (drylands vesper mouse); Csp, *Calomys* species; Hsap, *Homo sapiens* (human); Nalb, *Neotoma albigula* (white-throated woodrat); Ngui, *Neacomys guianae* (Guiana bristly mouse); Nmic, *Neotoma micropus* (southern plains woodrat); Oalb, *Oryzomys albigularis* (Tomes's oryzomys); Obic, *Oecomys bicolor* (bicolored arboreal rice rat); Obuc, *Oryzomys buccinatus* (Paraguayan rice rat); Ocap, *Oryzomys capito* (large-headed rice rat); Osp, *Oryzomys* species; Pcal, *Peromyscus californicus* (California mouse); Sals, *Sigmodon alstoni* (Alston's cotton rat); Shis, *Sigmodon hispidus* (hispid cotton rat). The LCMV strain WE is an Old World arenavirus and was included in the analyses to infer the ancestral node within the group of New World arenaviruses.

DISCUSSION

The present-day principal host relationships of some South American arenaviruses appear to be a consequence of a long-term shared evolutionary relationship between the *Arenaviridae* and the subfamily Sigmodontinae (Bowen *et al*., 1997). Evidence for this ancient virus-rodent host relationship includes the present-day association of phylogenetically closely related arenaviruses with phylogenetically closely related sigmodontine rodents, for example -- JUNV with the drylands vesper mouse (*C. musculinus*) and MACV with the large vesper mouse (*C. callosus*). The results of the analyses of the amino acid sequence data and the analyses of the nucleotide sequence data in this study indicate that CTNV is phylogenetically most closely related to WWAV. Together, the association of CTNV with the southern plains woodrat in southern Texas and WWAV with the white-throated woodrat in northwestern New Mexico is evidence that the present-day principal host relationships of some New World arenaviruses are a consequence of a long-term shared evolutionary relationship between the *Arenaviridae* and the subfamily Neotominae, which is exclusively North American (Musser and Carleton, 2005).

Previous studies revealed that some arenaviruses in association with their principal hosts are geographically widely distributed in nature. For example, GTOV in association with the short-tailed cane mouse (*Zygodontomys brevicauda*) and PIRV in association with Alston's cotton rat (*S. alstoni*) are distributed across a region that comprises 5 states on the plains of western Venezuela (Weaver *et al*., 2000, 2001). The present-day geographical range of the southern plains woodrat extends from northwestern New Mexico, southeastern Colorado, and southwestern Kansas through western Texas into northern Chihuahua, eastern San Luis Potosi, and southern Tamaulipas in Mexico (Musser and Carleton, 2005). Antibody to a Tacaribe complex virus previously was found in 2 (18.2%) of 11 southern plains woodrats captured in southeastern Colorado (Calisher *et al*., 2001), 6 (66.7%) of 9 southern plains woodrats captured in southeastern New Mexico, 2 (40%) of 5 southern woodrats captured in northwestern Texas, and 1

(33.3%) of 3 southern plains woodrats captured in western Texas (R. D. Bradley and M. L. Milazzo, unpublished data). Antibody to a Tacaribe complex virus also was found in 1 (16.7%) of 6 southern plains woodrats captured in Nuevo León and 1 (7.7%) of 13 southern plains woodrats captured in Tamaulipas (A. Barragán-Gómez, unpublished data). Thus, the geographical range of CTNV may occupy a large region within the geographical range of the southern plains woodrat.

Six arenaviral species naturally cause severe febrile disease in humans: GTOV, JUNV, MACV, SABV, LCMV and LASV (Peters, 2002). The human health significance of CTNV and other arenaviruses associated with neotomine or sigmodontine rodents in North America is the subject of ongoing research supported by the United States Public Health Service, National Institutes of Health.

References

Archer, A.M., Rico-Hesse, R., 2002. High genetic divergence and recombination in arenaviruses from the Americas. *Virology* 304, 274-281.

Armstrong, C., Lillie, R.D., 1934. Experimental lymphocytic choriomeningitis of monkeys and mice produced by a virus encountered in studies of the 1933 St Louis encephalitis epidemic. *Public Health Rep* 49, 1019-1027.

Auperin, D., Compans, R.W., Bishop, D.H.L., 1982a. Nucleotide sequence conservation at the 3' termini of the virion RNA species of New World and Old World arenaviruses. *Virology* 121, 200-203.

Auperin, D., Dimock, K., Cash, P., Rawls, W.E., Leung, W-C., Bishop, D.H.L., 1982b. Analyses of the genomes of prototype Pichinde arenavirus and a virulent derivative of Pichinde Munchique: evidence for sequence conservation at the 3' termini of their viral RNA species. *Virology* 116, 363-367.

Auperin, D., McCormick, J.B., 1989. Nucleotide sequence of the Lassa virus (Josiah strain) S genome RNA and amino acid sequence comparison of the N and GPC proteins to other arenaviruses. *Virology* 168, 421-425.

Barton, L.L., Peters, C.J., Ksiazek, T.G., 1995. Lymphocytic choriomeningitis virus: an unrecognized teratogenic pathogen. *Emerg Infect Dis* 1, 152-153.

Bausch, D.G., Rollin, P.E., Demby, A.H., Coulibaly, M., Kanu, J., Conteh, A.S., Wagoner, K.D., McMullan, L.K., Bowen, M.D., Peters, C.J., Ksiazek, T.G., 2000. Diagnosis and clinical virology of Lassa fever as evaluated by enzyme-linked immunosorbent assay, indirect fluorescent antibody test, and virus isolation. *J Clin Microbiol* 38, 2670-2677.

Bennett, S.G., Milazzo, M.L., Webb, J.P. Jr., Fulhorst, C.F., 2000. Arenavirus antibody in rodents indigenous to coastal southern California. *Am J Trop Med Hyg* 62, 626-630.

Beyer, W.R., Popplau, D., Garten, W., von Laer, D., Lenz, O., 2003. Endoproteolytic processing of the lymphocytic choriomenngitis virus glycoprotein by the subtilase SKI-1/S1P. *J Virol* 77, 2866-2872.

Bishop, D.H.L., 1986. Ambisense RNA genomes of arenaviruses and phleboviruses. *Adv Virus Res* 31, 1-51.

Biosafety in Microbiological and Biomedical Laboratories (BMBL) Fifth edition, 2007. Chosewood, L.C., Wilson, D.E. (Eds.) U.S. Government Printing Office, Washington, D.C.

Borden, K.L., 2000. RING domains: master builders of molecular scaffolds? *J Mol Biol* 295, 1103-1112.

Bowen, M.D., Peters, C.J., Nichol, S.T., 1996. The phylogeny of New World (Tacaribe complex) arenaviruses. *Virology* 219, 285-290.

Bowen, M.D., Peters, C.J., Nichol, S.T., 1997. Phylogenetic analysis of the *Arenaviridae*: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Mol Phylogenet Evol* 8, 301-316.

Bowen, M.D., Rollin, P.E., Ksiazek, T.G., Hustad, H.L., Bausch, D.G., Demby, A.H., Bajani, M.D., Peters, C.J., Nichol, S.T., 2000. Genetic diversity among Lassa virus strains. *J Virol* 74, 6992-7004.

Buchmeier, M.J., Bowen, M.D., Peters, C.J., 2001. *Arenaviridae*: the viruses and their replication. In: Knipe, D.M., Howley, P.M., Chanock, R.M., Melnick, J.L., Monath, T.P., Roizman, B., Straus, S.E. (Eds.), Fields Virology, Fourth edition, volume II. Lippincott, Williams & Wilkins, Philadelphia, pp. 1635-1668.

Buchmeier, M.J., 2002. Arenaviruses: protein structure and function. *Curr Top Microbiol Immunol* 262, 159-173.

Buckley, S.M., Casals, J., 1970. Lassa fever, a new virus disease of man from West Africa. III. Isolation and characterization of the virus. *Am J Trop Med Hyg* 19, 680-691.

Cajimat, M.N.B., Fulhorst, C.F., 2004. Phylogeny of the Venezuelan arenaviruses. *Virus Res* 102, 199-206.

Calisher, C.H., Tzianabos, T., Lord, R.D., Coleman, P.H., 1970. Tamiami virus, a new member of the Tacaribe group. *Am J Trop Med Hyg* 19, 520-526.

Calisher, C.H., Nabity, S., Root, J.J., Fulhorst, C.F., Beaty, B.J., 2001. Transmission of an arenavirus in white-throated woodrats (*Neotoma albigula*) in southeastern Colorado, 1995-1999. *Emerg Infect Dis* 7, 397-402.

Casals, J., 1965. Serological studies on Junin and Tacaribe viruses. *Am J Trop Med Hyg* 14, 794-796.

Casals, J., Buckley, S.M., Cedeno, R., 1975. Antigenic properties of the arenaviruses. *Bull World Health Org* 52, 421-427.

CDC, 1994. International Notes Bolivian hemorrhagic fever – El Beni Department, Bolivia, 1994. *MMWR Morbid Mortal Wkly Rep* 43, 943-946.

CDC, 2005. Lymphocytic choriomeningitis virus infection in organ transplant recipients – Massachusetts, Rhode Island, 2005. *MMWR Morbid Mortal Wkly Rep* 54, 537-539.

Charrel, R.N., de Lamballerie, X., Fulhorst, C.F., 2001. The Whitewater Arroyo virus: natural evidence for genetic recombination among Tacaribe serocomplex viruses (family *Arenaviridae*). *Virology* 283, 161-166.

Charrel, R.N., Feldmann, H., Fulhorst, C.F., Khelifa, R., de Chesse, R., de Lamballerie, X., 2002. Phylogeny of New World arenaviruses based on the complete coding sequences of the small genomic segment identified an evolutionary lineage produced by intrasegmental recombination. *Biochem Biophys Res Commun* 296, 1118-1124.

Charrel, R.N., Lemasson, J.J., Garbutt, M., Khelifa, R., de Micco, P., Feldmann, H., de Lamballerie, X., 2003. New insights into the evolutionary relationships between arenaviruses provided by comparative analysis of small and large segment sequences. *Virology* 317, 191-196.

Childs, J.E., Peters, C.J., 1993. Ecology and epidemiology of arenaviruses and their hosts. In: Salvato, M.S. (Ed.), The *Arenaviridae*. Plenum Press, New York, pp. 331-384.

Clegg, J.C., Wilson, S.M., Oram, J.D., 1991. Nucleotide sequence of the S RNA of Lassa virus (Nigerian strain) and comparative analysis of arenavirus gene products. *Virus Res* 18, 151-164.

Clegg, J.C.S., 2002. Molecular phylogeny of the arenaviruses. *Curr Top Microbiol Immunol* 262, 1-24.

Cornu, T.I., de la Torre, J.C., 2001. RING finger Z protein of Lymphocytic choriomeningitis virus (LCMV) inhibits transcription and replication of an LCMV S-segment minigenome. *J Virol* 75, 9415-9426.

Cornu, T.I., de la Torre, J.C., 2002. Characterization of the arenavirus RING finger Z protein regions required for Z-mediated inhibition of viral RNA synthesis. *J Virol* 76, 6678-6688.

Demby, A.H., Chamberlain, J., Brown, D.W., Clegg, .C.S., 1994. Early diagnosis of Lassa fever by reverse transcription-PCR. *J Clin Microbiol* 32, 2898-2903.

Djavani, M., Lukashevich, I.S., Sanchez, A., Nichol, S.T., Salvato, M.S., 1997. Completion of the Lassa fever virus sequence and identification of a RING finger open reading frame at the L RNA 5' end. *Virology* 235, 414-418.

Downs, W.G., Anderson, C.R., Spence, L., Aitken, T.H.G., Greenhall, A.H., 1963. Tacaribe virus, a new agent isolated from Artibeus bats and mosquitoes in Trinidad, West Indies. *Am J Trop Med Hyg* 12, 640-646.

Edwards, C., Bradley, R.D., 2002. Molecular systematics of the genus *Neotoma. Mol Phylogenet Evol* 25, 489-500.

Eichler, R., Lenz, O., Strecker, T., Eickmann, M., Klenk, H.-D., Garten, W., 2003. Signal peptide of Lassa virus glycoprotein GP-C exhibits an unusual length. *FEBS Lett* 538, 203-206.

Emonet, S., Lemasson, J.-J., Gonzalez, J.-P., de Lamballerie, X., Charrel, R.N., 2006. Phylogeny and evolution of Old World arenaviruses. *Virology* 350, 251-257.

Enria, D., Franco, S.G., Ambrosio, A., Vallejos, D., Levis, S., Maiztegui, J., 1986. Current status of the treatment of Argentine hemorrhagic fever. *Med Microbiol Immunol* 175, 173-176.

Enria, D.A., Maiztegui, J.I., 1994. Antiviral treatment of Argentine hemorrhagic fever. *Antiviral Res* 23, 23-31.

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

Frame, J.D., Baldwin, J.M. Jr., Gocke, D.J., Troup, J.M., 1970. Lassa fever, a new virus disease of man from West Africa. I. Clinical description and pathological findings. *Am J Trop Med Hyg* 19, 670-676.

Fulhorst, C.F., Bowen, M.D., Ksiazek, T.G., Rollin, P.E., Nichol, S.T., Kosoy, M.Y., Peters, C.J., 1996. Isolation and characterization of Whitewater Arroyo virus, a novel North American arenavirus. *Virology* 224, 114-120.

Fulhorst, C.F., Bowen, M.D., Salas, R.A., de Manzione, N.M.C., Duno, G., Utrera, A., Ksiazek, T.G., Peters, C.J., Nichol, S.T., de Miller, E., Tovar, D., Ramos, B., Vasquez, C., Tesh, R.B. 1997. Isolation and characterization of Pirital virus, a novel South American arenavirus. *Am J Trop Med Hyg* 56, 548-553.

Fulhorst, C.F., Bowen, M.D., Salas, R.A., Duno, G., Utrera, A., Ksiazek, T.G., de Manzione, N.M.C., de Miller, E., Vasquez, C., Peters, C.J., Tesh, R.B., 1999. Natural rodent host associations of Guanarito and Pirital viruses (family *Arenaviridae*) in central Venezuela. *Am J Trop Med Hyg* 61, 325-330.

Fulhorst, C.F., Charrel, R.N., Weaver, S.C., Ksiazek, T.G., Bradley, R.D., Milazzo, M.L., Tesh, R.B., Bowen, M.D., 2001. Geographical range and genetic diversity of

Whitewater Arroyo virus (family *Arenaviridae*) in the southwestern United States. *Emerg Infect Dis* 7, 403-407.

Fulhorst, C.F., Bennett, S.G., Milazzo, M.L., Murray, H.L., Webb, J.P. Jr., Cajimat, M.N.B., Bradley, R.D., 2002a. Bear Canyon virus: an arenavirus naturally associated with *Peromyscus californicus* (California mouse). *Emerg Infect Dis* 8, 717-720.

Fulhorst, C.F., Milazzo, M.L., Carroll, D.S., Charrel, R.N., Bradley, R.D., 2002b. Natural host relationships and genetic diversity of Whitewater Arroyo virus in Southern Texas. *Am J Trop Med Hyg* 67, 114-118.

Garcia, J.B., Morzunov, S.P., Lewis, S., Rowe, J., Calderon, G., Enria, D., Sabattini, M., Buchmeier, M.J., Bowen, M.D., St. Jeor, S.C., 2000. Genetic diversity of Junin virus in Argentina: geographic and temporal patterns. *Virology* 272, 127-136.

Gonzalez, J.P., McCormick, J.B., Saluzzo, J.F., Herve, J.P., Georges, A.J., Johnson, K.M., 1983. An arenavirus isolated from wild-caught rodents (*Praomys* species) in the Central African Republic. *Intervirology* 19, 105-112.

Gonzalez, J.P., Sanchez, A., Rico-Hesse, R., 1995. Molecular phylogeny of Guanarito virus, an emerging arenavirus affecting humans. *Am J Trop Med Hyg* 53**,** 1-6.

Gregg, M.B., 1975. Recent outbreaks of lymphocytic choriomeningitis in the United States of America. *Bull World Health Organ* 52, 549-553.

Günther, S., Emmerich, P., Laue, T., Kühle, O., Asper, M., Jung, A., Grewing, T., ter Meulen, J., Schmitz, H., 2000. Imported lassa fever in Germany: molecular characterization of a new Lassa virus strain. *Emerg Infect Dis* 6, 466-76.

Günther, S., Weisner, B., Roth, A., Grewing, T., Asper, M., Drosten, C., Emmerich, P., Petersen, J., Wilczek, M., Schmitz, H., 2001. Lassa fever encephalopathy: Lassa virus in cerebrospinal fluid but not in serum. *J Infect Dis* 184, 345-349.

Hall, E.R., Kelson, K.R., 1959. The mammals of North America. The Ronald Press Co., New York, pp. 547-1084.

Harrison, L.H., Halsey, N.A., McKee, K.T. Jr., Peters, C.J., Barrera Oro, J.G., Briggiler, A.M., Feuillade, M.R., Maiztegui, J.I., 1999. Clinical case definitions for Argentine hemorrhagic fever. *Clin Infect Dis* 28, 1091-1094.

Huelsenbeck, J.P., Ronquist, F.R., 2001. MRBAYES: Bayesian inference of phylogeny. *Biometrics* 17, 754-755.

Jacamo,R., Lopez, N., Wilda, M., Franze-fernandez, M.T., 2003. Tacaribe virus Z protein interacts with the L polymerase protein to inhibit viral RNA synthesis. *J Virol* 77, 10383-10393.

Jahrling, P.B., Peters, C.J. 1992. Lymphocytic choriomeningitis virus. *Arch Pathol Lab Med* 186, 486-488.

Jay, M.T., Glaser, C., Fulhorst, C.F., 2005. The arenaviruses. *J Am Vet Med Assoc* 227, 904-915.

Jennings, W.L., Lewis, A.L., Sather, G.E., Pierce, L.V., Bond, J.O., 1970. Tamiami virus in the Tampa Bay area. *Am J Trop Med Hyg* 19, 527-536.

Johnson, K.M., Wiebenga, N.H., Mackenzie, R.B., Kuns, M.L., Tauraso, N.M., Shelokov, A., Webb, P.A., Justines, G., Beye, H.K., 1965. Virus isolations from human cases of hemorrhagic fever in Bolivia. *Proc Soc Exp Biol Med* 118, 113-118.

Johnson, K.M., Kuns, M.L., Mackenzie, R.B., Webb, P.A., Yunker, C.E., 1966. Isolation of Machupo virus from wild rodent *Calomys callosus*. *Am J Trop Med Hyg 15:* 103--106.

Johnson, K.M., Taylor, P., Elliott, L.H., Tomori, O., 1981. Recovery of a Lassa-related arenavirus in Zimbabwe. *Am J Trop Med Hyg* 30, 1291-1293.

Kilgore, P.E., Peters, C.J., Mills, J.N., Rollin, P.E., Armstrong, L., Khan, A.S., Ksiazek, T.G., 1995. Prospects for the control of Bolivian hemorrhagic fever. *Emerg Infect Dis* 1, 97-100.

Kilgore, P.E., Ksiazek, T.G., Rollin, P.E., Mills, J.N., Villagra, M.R., Montenegro, M.J., Costales, M.A., Paredes, L.C., Peters, C.J., 1997. Treatment of Bolivian hemorrhagic fever with intravenous ribavirin. *Clin Infect Dis* 24, 718-722.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M., 2001. *MEGA*2: molecular evolutionary genetics analysis software. *Bioinformatics* 17, 1244-1245.

Kumar, S., Tamura, K., Nei, M., 2004. *MEGA*3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5, 150-63.

Lenz, O., ter Meulen, J., Klenk, H.-D., Seidah, N.G., Garten, W., 2001. The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci USA* 98, 12701-12705.

Lisieux, T., Coimbra, M., Nassar, E.S., Burattini, M.N., de Souza, L.T., Ferreira, I., Rocco, I.M., Travassos da Rosa, A.P., Vasconcelos, P.F., Pinheiro, F.P., LeDuc,

J.W., Rico-Hesse, R., Gonzalez, J.-P., Jahrling, P.B., Tesh, R.B., 1994. New arenavirus isolated in Brazil. *Lancet* 343, 391-392.

Lozano, M.E., Ghiringhelli, P.D., Romanowski, V., Grau, O., 1993. A simple nucleic acid amplification assay for the rapid detection of Junin virus in whole blood samples. *Virus Res* 27, 37-53.

Lozano, M.E., Enria, D., maiztegui, J.I., Grau, O., Romanowski, V., 1995. Rapid diagnosis of Argentine hemorrhagic fever by reverse transcriptase PCR-based assay. *J Clin Microbiol* 33, 1327-1332.

Lukashevich, I.S., 1992. Generation of reassortants between African arenaviruses. *Virology* 188, 600-605.

Mackenzie, R.B., Beye, H.K., Valverde, L., Garron, H., 1964. Epidemic hemorrhagic fever in Bolivia. I. A preliminary report of the epidemiology and clinical findings in a new epidemic area in South America. *Am J Trop Med Hyg* 13, 620-625.

Mackenzie, R.B., 1965. Epidemiology of Machupo virus infection. I. Pattern of human infection, San Joaquin, Bolivia, 1962-1964. *Am J Trop Med Hyg* 14, 808-813.

Maiztegui, J.I., 1975. Clinical and epidemiological patterns of Argentine hemorrhagic fever. *Bull World Health Org* 52, 567-575.

Maiztegui, J.I., Fernandez, N.J., de Damilano, A.J., 1979. Efficacy of immune plasma in treatment of Argentine hemorrhagic fever and association between treatment and a late neurological syndrome. *Lancet* 2, 1216-1217.

de Manzione, N., Salas, R.A., Paredes, H., Godoy, O., Rojas, L., Araoz, F., Fulhorst, C.F., Ksiazek, T.G., Mills, J.N., Ellis, B.A., Peters, C.J., Tesh, R.B., 1998. Venezuelan hemorrhagic fever: clinical and epidemiological studies of 165 cases. *Clin Infect Dis* 26, 308-13.

Matocq, M.D., 2002. Morphological and molecular analysis of a contact zone in the *Neotoma fuscipes* species complex. J Mammal 83, 866-883.

McCormick, J.B., Webb, P.A., Krebs, J.W., Johnson, K.M., Smith, E.S., 1987. A prospective study of the epidemiology and ecology of Lassa fever. *J Infect Dis* 155, 437-444.

McCormick, J.B., Fisher-Hoch, S.P., 2002. Lassa fever. *Curr Top Microbiol Immunol* 262, 75-109.

Mettler, N.E., Casals, J., Shope, R.E., 1963. Study of the antigenic relationships between the etiological agent of Argentinian hemorrhagic fever, Junin virus and other arthropod-borne viruses. *Am J Trop Hyg Hyg* 12, 647-652.

Meyer, B.J., de la Torre, J.C., Southern, P.J., 2002. Arenaviruses : genomic RNAs, transcription, and replication. *Curr Top Microbiol Immunol* 262, 139-157.

Mills, J.N., Ellis, B.A., McKee, K.T. Jr., Ksiazek, T.G., Barrera Oro, J.G., Maiztegui, J.I., Calderon, G.E., Peters, C.J., Childs, J.E., 1991. Junin virus activitity in rodents from endemic and nonendemic loci in central Argentina. *Am J Trop Med Hyg* 44, 589-597.

Mills, J.N., Ellis, B.A., McKee, K.T. Jr., Calderon, G.E., Maiztegui, J.I.,Nelson, G.O., Ksiazek, T.G., Peters, C.J., Childs, J.E., 1992. A longitudinal study of Junin virus activity in the rodent reservoir of Argentine hemorrhagic fever. *Am J Trop Med Hyg* 47, 749-763.

Mills, J.N., Barrera Oro, J.G., Bressler, D.S., Childs, J.E., Tesh, R.B., Smith, J.F., Enria, D.A., Geisbert, T.W., McKee, K.T. Jr., Bowen, M.D., Peters, C.J., Jahrling, P.B., 1996. Characterization of Oliveros virus, a new member of the Tacaribe complex (Arenaviridae: Arenavirus). *Am J Trop Med Hyg* 54, 399-404.

Mills, J.N, Childs, J.E., 1998. Ecologic studies of rodent reservoirs: their relevance for human health. *Emerg Infect Dis* 4, 529-537.

Moncayo, A.C., Hice, C.L., Watts, D.M., Travassos da Rosa, A.P.A., Guzman, H., Russell, K.L., Calampa, C., Gozalo, A., Popov, V.L., Weaver, S.C., Tesh, R.B., 2001. Allpahuayo virus: a newly recognized arenavirus (*Arenaviridae*) from arboreal rice rats (*Oecomys bicolor* and O*ecomys paricola*) in northeastern Peru. *Virology* 284, 277-286.

Murphy, F.A., Webb, P.A., Johnson, K.M., Whitfield, S.G., 1969. Morphological comparison of Machupo with lymphocytic choriomeningitis virus: basis for a new taxonomic group. *Am J Trop Med Hyg* 4, 535-541.

Murphy, F.A., 1975. Arenavirus taxonomy: a review. *Bull World Health Org* 52, 389-391.

Musser, G.G., Carleton, M.D., 2005. Superfamily Muroidea. In: Wilson, D.E., Reeder, D.M. (Eds), Mammal species of the world: A taxonomic and geographic reference. Third edition. Johns Hopkins University Press, Baltimore, pp. 894-1522.

Nichol, S.T., 2001. *Bunyaviridae*. In: Knipe, D.M., Howley, P.M., Chanock, R.M., Melnick, J.L., Monath, T.P., Roizman, B., Straus, S.E. (Eds.), Fields Virology, Fourth edition, volume II. Lippincott, Williams & Wilkins, Philadelphia, pp. 1603-1633.

Nichol, S.T., Beaty, B.J., Elliott, R.M., Goldbach, R., Plyusnin, A., Schmaljohn, C.S., Tesh, R.B., 2005. Family *Bunyaviridae*. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, pp. 695-716.

Parodi, A.S., Greenway, D.Y., Rugiero, H.R., Rivero, E., Frigerio, M.J., Mettler, N.E., Garzon, F., Boxaca, M., de Guerrero, L.B., Nota, N.R., 1958. Sobre la etiologia del brote epidemico en Junin. *Dia Med* 30, 2300-2302.

Perez, M., Craven, R.C., de la Torre, J.C., 2003. The small RING finger protein Z drives arenavirus budding: implications for antiviral strategies. *Proc Natl Acad Sci USA* 100, 12978-12983.

Peters, C.J., Webb, P.A., Johnson, K.M., 1973. Measurement of antibodies to Machupo virus by the indirect fluorescent technique. *Proc Soc Exp Biol Med* 142, 526-531.

Peters, C.J., Kuehne, R.W., Mercado, R.R., Le Bow, R.H., Spertzel, R.O., Webb, P.A., 1974. Hemorrhagic fever in Cochabamba, Bolivia, 1971. *Am J Epidemiol* 99, 425-433.

Peters, C.J., 2002. Human infection with arenaviruses in the Americas. *Curr Top Microbiol Immunol 262:* 65--74.

Pinheiro, F.P., Shope, R.E., de Andrade, A.H.P., Bensabath, G., Cacios, G.V., Casals, J., 1966. Amapari, a new virus of the Tacaribe group from rodents and mites of Amapa territory, Brazil. *Proc Soc Exp Biol Med* 122, 531-535.

van Regenmortel, M.H.V., 2000. Introduction to the species concept in virus taxonomy. In: van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M. (Eds.) Virus Taxonomy: Seventh Report of the International Committee on Taxonomy of Viruses. Academic Press, San Diego, CA, pp. 3-16.

Riera, L.M., Feuillade, M.R., Saavedra, M.C., Ambrosio, A.M.,1997. Evaluation of an enzyme immunosorbent assay for the diagnosis of Argentine haemorrhagic fever. *Acta Virol* 41, 305-310.

Rivers, T.M., Scott, T.F.M., 1935. Meningitis in man caused by a filterable virus. *Science* 81, 439-440.

Riviere, Y., Oldstone, M.B., 1986. Genetic reassortants of lymphocytic choriomeningitis virus: unexpected disease and mechanism of pathogenesis. *J Virol* 59, 363-368.

Rowe, W.P., Murphy, F.A., Bergold, G.H., Casals, J., Hotchin, J., Johnson, K.M., Lehmann-Grube, F., Mims, C.A., Traub, E., 1970a. Arenoviruses: proposed name for a newly defined virus group. *J Virol* 5, 651-652.

Rowe, W.P., Pugh, W.E., Webb, P.A., Peters, C.J., 1970b. Serological relationships of the Tacaribe complex of viruses to lymphocytic choriomeningitis virus. *J Virol* 5, 289-292.

Salas, R.,Rico-Hesse, R., Godoy, O., Ramos, B., Jaime, E., Manzione, N. de., Shope, R.E., Bruzual, R., Taibo, M.E., Vasquez, C., Querales, J., Tesh, R.B., Betancourt, A., Pacheco, M.E., Tamayo, J.G., Araoz, F., Querales, J., 1991. Venezuelan haemorrhagic fever. *Lancet* 338, 1033-1036.

Salazar-Bravo, J., Ruedas, L.A., Yates, T.L., 2002. Mammalian reservoirs of arenaviruses. *Curr Top Microbiol Immunol* 262, 26-63.

Salvato, M.S., Clegg, J.C.S., Buchmeier, M.J., Charrel, R.N., Gonzalez, J.P., Lukashevich, I.S., Peters, C.J., Rico-Hesse, R., Romanowski, V., 2005. Family *Arenaviridae*. In: Fauquet, C.M., Mayo, M.A., Maniloff, J., Desselberger, U., Ball, L.A. (Eds.), Virus Taxonomy: Eighth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press, pp. 725-733.

Sanchez, A., Pifat, D.Y., Kenyon, R.H., Peters, C.J., McCormick, J.B., Kiley, M.P., 1989. Junin virus monoclonal antibodies: characterization and cross-reactivity with other arenaviruses. *J Gen Virol* 70, 1125-1132.

Schmitz, H., Köhler, B., Laue, T., Drosten, C., Veldkamp, P.J., Günther, S., Emmerich, P., Geisen, H.P., Fleischer, K., Beersma, M.F.C., Hoerauf, A., 2002. Monitoring of clinical and laboratory data in two cases of imported Lassa fever. *Microbes Infect* 4, 43-50.

Southern, P.J., 1996. Arenaviruses. In: Fields, B.N., Knipe, D.M., Howley, P.M.(Eds.), Fields Virology, 3rd edition, volume II. Lippincott-Raven, Philadelphia, pp. 1505-1520.

Steppan, S., Adkins, R., Anderson, J., 1994. Phylogeny and divergence-date estimates of rapid radiations in muroid rodents based on multiple nuclear genes. *Syst Biol* 53, 533-553.

Stinebaugh, B.J., Schloeder, F.X., Johnson, K.M., macKenzie, R.B., Entwisle, G., De Alba, E., 1966. Bolivian hemorrhagic fever a report of four cases. *Am J Trop Med Hyg* 40, 217-230.

Strecker, T., Eichler, R., ter Meulen, J., Weissenhorn, W., Klenk, H.-D., Garten, W., Lenz, O., 2003. Lassa virus Z protein is a matrix protein and sufficient for the release of virus-like particles. *J Virol* 77, 10700-10705.

Swanepoel, R., Leman, P.A., Shepherd, A.J., Shepherd, S.P., Kiley, M.P., McCormick, J.B., 1985. Identification of Ippy as a Lassa-fever-related virus. *Lancet* 8429, 639.

Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4. Sinauer Associates, Sunderland, Massachusetts.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W (1.7): improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choices. *Nucleic Acids Res* 22, 4673-4680.

Tesh, R.B., Jahrling, P.B., Salas, R., Shope, R.E., 1994. Description of Guanarito virus (Arenaviridae: *Arenavirus*), the etiologic agent of Venezuelan hemorrhagic fever. *Am J Trop Med Hyg* 50, 452-459.

Traub, E.A. 1935. A filterable virus recovered from white mice. *Science* 81, 298-299.

Vieth, S., Torda, A.E., Asper, M., Schmitz, H., Günther, S., 2004. Sequence analysis of L RNA of Lassa virus. *Virology* 318, 153-168.

Vieth, S., Drosten, C., Charrel, R., Feldmann, H., Günther, S., 2005. Establishment of conventional and fluorescence resonance energy transfer-based real-time PCR assays for detection of pathogenic New World arenaviruses. *J Clin Virol* 32, 229-235.

Weaver, S.C., Salas, R.A., de Manzione, N., Fulhorst, C.F., Duno, G., Utrera, A., Mills, J.N., Ksiazek, T.G., Tovar, D., Tesh, R.B., 2000. Guanarito virus (*Arenaviridae*) isolates from endemic and outlying localities in Venezuela: sequence comparisons among and within strains isolated from Venezuelan hemorrhagic fever patients and rodents. *Virology* 266, 189-195.

Weaver, S.C., Salas, R.A., de Manzione, N., Fulhorst, C.F., Travassos da Rosa, A.P.A., Duno, G., Utrera, A., Mills, J.N., Ksiazek, T.G., Tovar, D., Guzman, H., Kang, W.,

Tesh, R.B., 2001. Extreme genetic diversity among Pirital virus (*Arenaviridae*) isolates from western Venezuela. *Virology* 285, 110-118.

Webb, P.A., Johnson, K.M., MacKenzie, R.B., 1969. The measurement of specific antibodies in Bolivian hemorrhagic fever by neutralization of virus plaques. *Proc Soc Exp Biol Med* 130, 1013-1019.

WHO, 1995. Re-emergence of Bolivian hemorrhagic fever. *Wkly Epid Rec* 3, 16-17.

Wulff, H., McIntosh, B.M., Hamner, D.B., Johnson, K.M., 1977. Isolation of an arenavirus closely related to Lassa virus from *Mastomys natalensis* in south-east Africa. *Bull World Health Org* 155, 441-444.

Vita

Maria Natividad B. Cajimat was born on January 6, 1976 to Mariano and Natividad Cajimat of Cagayan, Republic of the Philippines. She finished high school in 1993 and then attended the College of Science at the University of the Philippines in Quezon City. Maria received her Bachelor's degree in Molecular Biology and Biotechnology (*cum laude*) in April 1997. She worked as a research associate at the Research Institute for Tropical Medicine in Alabang in the Republic of the Philippines from 1997 through 2000 and began her graduate studies at the University of Texas Medical Branch in August 2000. She joined the laboratory of Dr. Charles F. Fulhorst in the Fall of 2001.

Prior to completing the writing of her dissertation, Maria joined the laboratory of Dr. Sonja R. Gerrard at the School of Public Health of the University of Michigan in Ann Arbor. She currently is working as a post-doctoral level scientist on a research project that involves collaboration between Dr. Gerrard's laboratory and Dr. Fulhorst's laboratory. This project entails rescue of infectious hantavirus entirely from cloned cDNA.

Education

B.S., April 1997, University of the Philippines, Diliman, Quezon City, Philippines

Publications

Cajimat, M.N.B., Milazzo, M.L., Hess, B.D., Rood, M.P., Fulhorst, C.F., 2007. Principal host relationships and evolutionary history of the North American arenaviruses. *Virology* (2007 Jul 10, Epub ahead of print).

Cajimat, M.N.B., Milazzo, M.L., Bradley, R.D., Fulhorst, C.F., 2007. *Catarina virus*, an arenaviral species principally associated with *Neotoma micropus* (southern plains woodrat) in Texas. *Am J Trop Med Hyg* (In press, June 2007).

Milazzo, M.L., **Cajimat, M.N.B.**, Hanson, J.D., Bradley, R.D., Quintana, M., Sherman, C., Velasquez, R.T., Fulhorst, C.F., 2006. *Catacamas virus*, a hantaviral species naturally associated with *Oryzomys couesi* (Coues' oryzomys) in Honduras. *Am J Trop Med Hyg* 75, 1003-1010.

Fulhorst, C.F., **Cajimat, M.N.B.**, Utrera, A., Milazzo, M.L., Duno, G.M., 2004. Maporal virus, a hantavirus associated with the fulvous pygmy rice rat (*Oligoryzomys fulvescens*) in western Venezuela. *Virus Res* 104, 139-144.

Cajimat, M.N.B., Fulhorst, C.F., 2004. Phylogeny of the Venezuelan arenaviruses. *Virus Res* 102,199-206.

Newman, M., Suk, F.-M., **Cajimat, M.**, Chua, P.K., Shih, C., 2003. Stability and morphology comparisons of self-assembled virus-like particles from wild-type and mutant human hepatitis B virus capsid proteins. *J Virol* 77, 12950-12960.

Fulhorst, C.F., Bennett, S.G., Milazzo, M.L., Murray, H.L. Jr., Webb, J.P. Jr, **Cajimat, M.N.**, Bradley, R.D., 2002. Bear Canyon virus, an arenavirus naturally associated with *Peromyscus californicus* (California mouse). *Emerg Infect Dis* 8, 717-721.

Santiago, M.L.O., Santiago, E.G., Hafalla, J.C.R., Manalo, M.A., Orantia, L., **Cajimat, M.N.B.**, Martin, C., Cuaresma, C., Dominguez, C., Borromeo, M.E., DeGroot, A.S., Flanigan, T.P., Carpenter, C.J., Mayer, K., Ramirez, B.L., 1998. Molecular epidemiology of HIV-1 infection in the Philippines, 1985-1997: transmission of subtypes B and E and potential emergence of subtypes C and F. *J Acquir Immune Defic Syndr Hum Retrovirol* 18, 260-269.