White Paper Application

Project Title: Complete genome sequencing of a collection of eastern equine encephalitis viruses (EEEV).

Authors: Albert J. Auguste, Robert B. Tesh and Scott C. Weaver.

Primary Investigator Contact:

Name	Scott C. Weaver
Position	Professor
Institution	University of Texas Medical Branch (UTMB)
Address	Department of Pathology, 301 University Blvd., Galveston
State	Texas
ZIP Code	77555-0609
Telephone	409 747 0758
Fax	409-747-2429
E-Mail	sweaver@utmb.edu

1. Executive Summary *(Please limit to 500 words.)*

Eastern equine encephalitis virus (EEEV) is a mosquito-borne alphavirus (family, *Togaviridae*) of significant public and veterinary health importance throughout North, Central and South America. It has a single-stranded positive sense RNA genome of about 11,700 nt in length. EEEV causes a very debilitating disease in humans with case-fatality rates ranging between 50 and 75%, and for equids, rates can achieve rates as high as 80%. It is currently classified as a select agent on the NIAID category B list of priority pathogens, and was being developed for use as a potential bioweapon during the cold war. Despite hundreds of isolations and the significant public and veterinary health problem this virus poses, the majority of EEEV genetic studies primarily involved only partial gene sequences of the structural or non-structural genes and only 10 complete genomes have been determined and deposited on GenBank to date. In contrast to EEEV, extensive genetic studies have been performed on related alphaviruses, such as Venezuelan equine encephalitis and Chikungunya viruses, which has lead to a detailed understanding of the mutations involved in emergence, genetic determinants of virulence and disease, and thorough characterization based on genealogy. In the absence of sufficient complete genome sequence data for EEEV, comparable studies cannot be performed thoroughly, and those few studies that relied on partial genomic fragments have already been very insightful in describing EEEV genetic diversity (Arrigo et al., 2010; Brault et al., 1999),

positively selected amino acid sites (Arrigo et al., 2010), and mutations associated with certain phenotypes (Cooper and Scott, 2001; Weaver et al., 1999). The proposed project aims to fill this gap in sequence data by elucidating the complete genomes for 100 EEEV strains that represent the entire geographic and temporal distribution of EEEV since its first isolation in the new world in 1933.

The resultant genetic data generated in this study would allow researchers to (i) characterize EEEV genetic diversity and describe its molecular epidemiology; (ii) identify genetic determinants of EEEV emergence and virulence; (iii) infer the demographic history and evolutionary dynamics of EEEV; (iv) infer the evolution of virulence and pathogenesis throughout EEEV's transmission history; (v) improve the ability to derive attenuated EEEV stains for vaccine development, (vi) generate infectious cDNA clones, and lastly; (vii) identify genetic factors underlying phenotypic differences among EEEV strains in the New World.

This project is proposed as a collaborative study between the NIAID Genome Sequencing Center, the University of Texas Medical Branch, The Connecticut Agricultural Experiment Station, South Florida University, Colorado State University, and the Division of Vector-Borne Diseases, and the Center for Disease Prevention and Control. We will ensure the generation of the best possible data set so further research in the aforementioned areas may be pursued without limitations. The data generated will also be useful for related alphavirus studies.

2. Justification

Eastern equine encephalitis virus (EEEV) is a single-strand positive sense RNA virus that belongs to the family, *Togaviridae*, genus *Alphavirus* (Morris, 1988). Its genome is ~11,700 nt long and is flanked by a 5' cap and 3' poly[A] tail. EEEV is classified as the only species in its antigenic complex (Calisher and Karabatsos, 1988; Calisher et al., 1980), but can be further classified into four (4) lineages or subtypes based on serologic and phylogenetic analyses of genomic sequences (Brault et al., 1999; Regenmortel et al., 2000). Lineage I primarily consists of North American and Caribbean isolates, and lineages II-IV consists of isolates from Central and South America (Brault et al., 1999). The majority of phylogenetic analyses that have been used to further characterize EEEV are based on short fragments of the structural genes (Arrigo, Adams, and Weaver, 2010; Brault et al., 1999) and this may be inadequate to accurately resolve the plausible variation among strains. To date only 10 EEEV genomes exist on GenBank, despite hundreds of isolations throughout North, Central and South America.

EEEV is considered to be among the most important of the encephalitic alphaviruses in the Americas and among the most virulent human viral pathogens. EEEV continues to be an important public health and veterinary concern with up to 21 human cases/year being reported in North America (CDC, 2010), and a recent equine outbreak occurred in Brazil, affecting 229 equids with a case fatality rate of 73% (Silva et al., 2011). EEE disease in humans ranges from relatively mild symptoms (such as fever, chills, malaise, and myalgias) to severe and often fatal encephalitis (Deresiewicz et al., 1997; Fothergill et al., 1938). Mortality after infection with EEEV has been estimated at 50 to 75% (which is by far the highest rate among the NIAID Category B encephalitic viruses), and recovering patients usually suffer with disabling and progressive mental and physical sequelae (Ayres and Feemster, 1949; Deresiewicz et al., 1997). Other clinical manifestations include personality disorders, seizures, and spastic paralysis (Ayres and Feemster, 1949; Calisher, 1994; Deresiewicz et al., 1997).

EEEV transmission cycles and epidemiological profiles vary widely between North and South American EEEV isolates. NA EEEV is transmitted in an enzootic cycle involving primarily avian vertebrate hosts such as passerine birds and the ornithophilic

vector *Culiseta melanura* (Morris, 1988; Scott and Weaver, 1989). However, since this vector is extremely ornithophilic and rarely bites mammals, bridge vectors are necessary to transmit the virus to other vertebrates including humans (Scott and Weaver, 1989). Many studies have been performed to determine the mechanism of EEEV overwintering in North America, and these include experimental studies that investigate the capacity of vector and/or host to perpetuate the virus over winter periods (Burkett-Cadena et al., 2011; Owen et al., 2011), as well as phylogenetic studies based on partial genomic data (Armstrong et al., 2008; Young et al., 2008). These phylogenetic studies were not only limited by the partial genomes used, but did contain sequence data from many of the Southern states where EEEV is endemic and is transmitted yearly. This study will provide the necessary sequence data to completely elucidate the phylogeographic history of EEEV, not only in North America, but also throughout new world. In South and Central America, epidemics of human disease are very uncommon but there are sporadic outbreaks in humans (few cases) (Aguilar et al., 2007; Corniou et al., 1972) and equids (involving thousands of cases) (Dietz et al., 1980; Sabattini et al., 1991). In Central and South America, mosquitoes of the subgenus *Culex Melanoconion* serve as the main enzootic vectors (Kondig et al., 2007; Turell et al., 2008; Turell et al., 2005; Walder et al., 1984), but in contrast to North American EEEV, a large number of small mammals and birds are thought to act as the main reservoirs/amplification hosts (de Souza-Lopes and de Abreu-Sacchetta, 1974; Scott and Weaver, 1989; Shope et al., 1966; Walder et al., 1984). Previous work suggests that humans are exposed to but do not develop apparent infection or severe disease with South American EEEV because of poor infectivity and/or avirulence of South American strains (Aguilar et al., 2007). Previously, this pattern was also observed in Panama. However, a 2010 outbreak of human encephalitis in Panama (including children) resulted in the diagnosis of several cases of EEE. Preliminary phylogenetic analyses with partial genomic sequences indicate that the etiologic strain is closely related to previous Panamanian EEEV isolates, suggesting a change in human virulence. Complete genomic sequences and carative studies with this 2010 isolate and those previously isolated in Panama and other regions of South America and can provide valuable insight into mutations associated with emergence or disease in humans.

EEEV is also a well-developed biological weapon and is very infectious via the aerosol route, and is currently classified as a select agent on the NIAID category B list of priority

pathogens (http://www.niaid.nih.gov/topics/biodefenserelated/biodefense/pages/cata.aspx). Despite its importance, EEEV has been relatively understudied, particularly when compared to other important alphaviruses such as Venezuelan equine encephalitis (VEEV) and Chikungunya viruses (CHIKV). Phylogenetic studies of genomic sequence data on VEEV (Anishchenko et al., 2006; Brault et al., 2002) and CHIKV (Tsetsarkin et al., 2011; Tsetsarkin et al., 2007) have all lead to very important insights and a better understanding of mutations (i.e. that arose in their genomes that have significantly increased their fitness, and has been associated with major epizootics and epidemics). Given the limited sequence data available, no extensive analyses to elucidate adaptive mutations have been performed but a recent study demonstrated the existence of positively selected amino acid sites in the structural genes (Arrigo et al., 2010), although it was suggested that this is not a major driving force in EEEV evolution. Cell specific (single-host-cell and alternating-host-cell) adaptive (phenotypic) changes of EEEV have also been demonstrated in both *in vitro* and *in vivo* studies (Cooper and Scott, 2001; Weaver et al., 1999). Although partial genomic sequencing clearly demonstrated synonymous and nonsynonymous substitutions associated with single-host-cell and alternating-host-cell passaging (Cooper and Scott, 2001; Weaver et al., 1999), due to the unavailability of complete genome sequence data, elucidating other possible mutations that me be associated were not possible.

These studies and the severe lack of EEEV genomic sequence data clearly underscore the importance of determining the complete genomic sequences of EEEV strains for understanding the possible mutations associated with epidemic emergence, EEEV pathogenesis, virulence, evolution and taxonomy.

Researchers interested in EEEV replication, pathogenesis, epidemiology, evolution (and related alphavirus evolution), diagnostics and detection for biodefense, and the development of a vaccine, will be able to use the genomic sequence information derived in this study to; (i) identify potential determinants of virulence that can be tested using reverse genetic approaches, (ii) understand the molecular epidemiology and evolutionary dynamics of EEEV using Bayesian coalescent analyses, (iii) describe the evolution of virulence and pathogenesis throughout EEEV's transmission history, (iv) accurately resolve the currently circulating genetic diversity through phylogenetic analyses, (v) vaccine related studies, (vi) studies based on infectious clone technology, and lastly (vii)

make better associations about the differences in genetic data and phenotypes among the different EEEV strains in the Americas. The public health community will also be able to use the EEEV database to determine the epidemic potential of newly arising strains or intentional introductions. The addition of complete genome sequences for representative isolates will therefore greatly improve the power of the sequence database for generating predictive data.

References cited

1. Aguilar, P. V., Robich, R. M., Turell, M. J., O'Guinn, M. L., Klein, T. A., Huaman, A., Guevara, C., Rios, Z., Tesh, R. B., Watts, D. M., Olson, J., and Weaver, S. C. (2007). Endemic eastern equine encephalitis in the Amazon region of Peru. *Am J Trop Med Hyg* 76(2), 293-8.
2. Anishchenko, M., Bowen, R. A., Paessler, S., Austgen, L., Greene, I. P., and Weaver, S. C. (2006). Venezuelan encephalitis emergence mediated by a phylogenetically predicted viral mutation. *Proc Natl Acad Sci U S A* 103(13), 4994-9.
3. Armstrong, P. M., Andreadis, T. G., Anderson, J. F., Stull, J. W., and Mores, C. N. (2008). Tracking eastern equine encephalitis virus perpetuation in the northeastern United States by phylogenetic analysis. *Am J Trop Med Hyg* 79(2), 291-6.
4. Arrigo, N. C., Adams, A. P., and Weaver, S. C. (2010). Evolutionary patterns of eastern equine encephalitis virus in North versus South America suggest ecological differences and taxonomic revision. *J Virol* 84(2), 1014-25.
5. Ayres, J. C., and Feemster, R. F. (1949). The sequelae of eastern equine encephalomyelitis. *N Engl J Med* 240(24), 960-2.
6. Brault, A. C., Powers, A. M., Chavez, C. L., Lopez, R. N., Cachon, M. F., Gutierrez, L. F., Kang, W., Tesh, R. B., Shope, R. E., and Weaver, S. C. (1999). Genetic and antigenic diversity among eastern equine encephalitis viruses from North, Central, and South America. *Am J Trop Med Hyg* 61(4), 579-86.
7. Brault, A. C., Powers, A. M., Holmes, E. C., Woelk, C. H., and Weaver, S. C. (2002). Positively charged amino acid substitutions in the e2 envelope glycoprotein are associated with the emergence of venezuelan equine encephalitis virus. *J Virol* 76(4), 1718-30.
8. Burkett-Cadena, N. D., White, G. S., Eubanks, M. D., and Unnasch, T. R. (2011). Winter biology of wetland mosquitoes at a focus of eastern equine encephalomyelitis virus transmission in Alabama, USA. *J Med Entomol* 48(5), 967-73.
9. Calisher, C. H. (1994). Medically important arboviruses of the United States and Canada. *Clin Microbiol Rev* 7(1), 89-116.
10. Calisher, C. H., and Karabatsos, N. (1988). Arbovirus serogroups: definition and geographic distribution. *The arboviruses: epidemiology and ecology* 1, 19-57.
11. Calisher, C. H., Shope, R. E., Brandt, W., Casals, J., Karabatsos, N., Murphy, F. A., Tesh, R. B., and Wiebe, M. E. (1980). Proposed antigenic classification of registered arboviruses I. Togaviridae. Alphavirus. *Intervirology* 14(5-6), 229-32.
12. CDC (May 2010). Centers for Disease Control and Prevention. Confirmed and probable eastern equine encephalitis cases, human,
13. United States, 1964–2008, by state. Available at
14. www.dev.cdc.gov/ncidod/dvbid/arbor/pdf/EEE_DOC.pdf.

15. Cooper, L. A., and Scott, T. W. (2001). Differential evolution of eastern equine encephalitis virus populations in response to host cell type. *Genetics* 157(4), 1403-12.
16. Corniou, B., Ardoin, P., Bartholomew, C., Ince, W., and Massiah, V. (1972). First isolation of a South American strain of Eastern Equine virus from a case of encephalitis in Trinidad. *Trop Geogr Med* 24(2), 162-7.
17. de Souza Lopes, O., and de Abreu Sacchetta, L. (1974). Epidemiological studies on eastern equine encephalitis virus in Sao Paulo, Brazil. *Rev Inst Med Trop Sao Paulo* 16(5), 253-8.
18. Deresiewicz, R. L., Thaler, S. J., Hsu, L., and Zamani, A. A. (1997). Clinical and neuroradiographic manifestations of eastern equine encephalitis. *N Engl J Med* 336(26), 1867-74.
19. Dietz, W. H., Jr., Galindo, P., and Johnson, K. M. (1980). Eastern equine encephalomyelitis in Panama: the epidemiology of the 1973 epizootic. *Am J Trop Med Hyg* 29(1), 133-40.
20. Fothergill, L. D., Dingle, J. H., Farber, S., and Connerley, M. L. (1938). Human encephalitis caused by the virus of the eastern variety of equine encephalomyelits. *N Eng J Med* 219, 411-22.
21. Kondig, J. P., Turell, M. J., Lee, J. S., O'Guinn, M. L., and Wasieloski, L. P., Jr. (2007). Genetic analysis of South American eastern equine encephalomyelitis viruses isolated from mosquitoes collected in the Amazon Basin region of Peru. *Am J Trop Med Hyg* 76(3), 408-16.
22. Morris, C. (1988). Eastern equine encephalitis. *In* "The arboviruses: epidemiology and ecology" (T. P. Monath, Ed.), pp. 2-20. CRC Press, Boca Raton.
23. Owen, J. C., Moore, F. R., Williams, A. J., Stark, L., Miller, E. A., Morley, V. J., Krohn, A. R., and Garvin, M. C. (2011). Test of recrudescence hypothesis for overwintering of eastern equine encephalomyelitis virus in gray catbirds. *J Med Entomol* 48(4), 896-903.
24. Regenmortel, M. H. V., Fauquet, C. M., Bishop, D. H. L., Carstens, E. B., Estes, M. K., Lemon, S. M., Maniloff, J., Mayo, M. A., McGeoch, D. J., and Pringle, C. R. (2000). "Virus taxonomy: classification and nomenclature of viruses. Seventh report of the International Committee on Taxonomy of Viruses." Academic Press.
25. Sabattini, M. S., Daffner, J. F., Monath, T. P., Bianchi, T. I., Cropp, C. B., Mitchell, C. J., and Aviles, G. (1991). Localized eastern equine encephalitis in Santiago del Estero Province, Argentina, without human infection. *Medicina (B Aires)* 51(1), 3-8.
26. Scott, T. W., and Weaver, S. C. (1989). Eastern equine encephalomyelitis virus: epidemiology and evolution of mosquito transmission. *Adv Virus Res* 37, 277-328.
27. Shope, R. E., de Andrade, A. H., Bensabath, G., Causey, O. R., and Humphrey, P. S. (1966). The epidemiology of EEE WEE, SLE and Turlock viruses, with special reference to birds, in a tropical rain forest near Belem, Brazil. *Am J Epidemiol* 84(3), 467-77.
28. Silva, M. L., Galiza, G. J., Dantas, A. F., Oliveira, R. N., Iamamoto, K., Achkar, S. M., and Riet-Correa, F. (2011). Outbreaks of Eastern equine encephalitis in northeastern Brazil. *J Vet Diagn Invest* 23(3), 570-5.
29. Tsetsarkin, K. A., Chen, R., Leal, G., Forrester, N., Higgs, S., Huang, J., and Weaver, S. C. (2011). Chikungunya virus emergence is constrained in Asia by lineage-specific adaptive landscapes. *Proc Natl Acad Sci U S A* 108(19), 7872-7.
30. Tsetsarkin, K. A., Vanlandingham, D. L., McGee, C. E., and Higgs, S. (2007). A single mutation in chikungunya virus affects vector specificity and epidemic potential. *PLoS Pathog* 3(12), e201.

31. Turell, M. J., O'Guinn, M. L., Dohm, D., Zyzak, M., Watts, D., Fernandez, R., Calampa, C., Klein, T. A., and Jones, J. W. (2008). Susceptibility of Peruvian mosquitoes to eastern equine encephalitis virus. *J Med Entomol* 45(4), 720-5.
32. Turell, M. J., O'Guinn, M. L., Jones, J. W., Sardelis, M. R., Dohm, D. J., Watts, D. M., Fernandez, R., Travassos da Rosa, A., Guzman, H., Tesh, R., Rossi, C. A., Ludwig, V., Mangiafico, J. A., Kondig, J., Wasieloski, L. P., Jr., Pecor, J., Zyzak, M., Schoeler, G., Mores, C. N., Calampa, C., Lee, J. S., and Klein, T. A. (2005). Isolation of viruses from mosquitoes (Diptera: Culicidae) collected in the Amazon Basin region of Peru. *J Med Entomol* 42(5), 891-8.
33. Walder, R., Suarez, O. M., and Calisher, C. H. (1984). Arbovirus studies in the Guajira region of Venezuela: activities of eastern equine encephalitis and Venezuelan equine encephalitis viruses during an interepizootic period. *Am J Trop Med Hyg* 33(4), 699-707.
34. Weaver, S. C., Brault, A. C., Kang, W., and Holland, J. J. (1999). Genetic and fitness changes accompanying adaptation of an arbovirus to vertebrate and invertebrate cells. *J Virol* 73(5), 4316-26.
35. Young, D. S., Kramer, L. D., Maffei, J. G., Dusek, R. J., Backenson, P. B., Mores, C. N., Bernard, K. A., and Ebel, G. D. (2008). Molecular epidemiology of eastern equine encephalitis virus, New York. *Emerg Infect Dis* 14(3), 454-60.

3. Rationale for Strain Selection

Dr. Scott Weaver (principal investigator) currently has the largest collection of EEEV strains at his institution. He has continuously collected EEEV strains throughout his research career and now has >200 strains that will serve as the primary source for this study. Dr. Robert B. Tesh (collaborator) is the director of the World Reference Center Emerging Viruses and Arboviruses (WRCEVA; see description below) and also has a very large collection of EEEV strains. Isolates will be sourced from both collections, and will represent all the geographic locations (North, Central and South America) for which EEEV has been isolated. Additionally, isolates representing the entire temporal period from its first isolation to present will be provided by several collaborators from Connecticut, Tennessee, Florida and the CDC (see letters of support) and will be incorporated in this study. Important isolates that exist and are not represented in our collection will be sought and included in this study. More specifically, isolates will be included that were isolated from different sources (i.e. humans, different mosquito species, and different avian and rodent species), from all countries where EEEV has been reported (also representing recorded epidemics and epizootics), and throughout the temporal distribution since its first isolation in 1933. To accurately represent EEEVs evolutionary history, we propose to

sequence 100 EEEV complete genomes that will fully represent the spatial and temporal distribution of EEEV.

4a. Approach to Data Production: Data Generation

This study would produce complete genome sequences for 100 retrospectively collected EEEV strains that represent the complete spatial and temporal distribution of EEEV's isolation history in the Americas. Since there is very little complete genome sequence data available for EEEV, these data will serve as reference genomes for EEEV strains throughout the Americas, serve as a readily available source of information for generating clones and identifying mutations (SNPs). cDNA will be generated from RNA extracted from semi-purified virus produced in Vero cell cultures, sheared, randomly amplified and sequenced using the 454 and illumina platforms at JCVI. These methods proved highly efficient for VEEV strains sequenced previously in this UTMB-JCVI collaboration. In addition, to exploit the availability of unpassaged EEEV strains, we will perform RT-PCR and deep sequence amplicons to gain insights into the within-host sequence variability and distribution of snPs.

4b. Approach to Data Production: Data Analysis

Researchers will be able to use the genetic data generated and associated metadata gain new insights and test hypotheses about EEEV pathogenesis, molecular determinants of virulence and emergence, phylogeography, evolution and ecology of the viruses under study. More specifically, the genomic sequence data generated here would facilitate (i) identifying potential determinants of emergence and virulence that can be tested using reverse genetic approaches, (ii) a better understanding of the molecular epidemiology and evolutionary dynamics of EEEV using Bayesian coalescent analyses, (iii) describing the evolution of virulence and pathogenesis throughout EEEV's transmission history, (iv) accurately resolving the currently circulating genetic diversity through phylogenetic analyses, (v) vaccinology studies, (vi) studies based on infectious clone technology, and lastly (vi) make better associations about the differences in genetic data and phenotypes among the different EEEV strains in the Americas.

5. Community Support and Collaborator Roles:

The following is a partial listing of US researchers with a major interest in this study.

1. Dr. Nicole Arrigo, Center for Infection and Immunity, Columbia University, New York.
2. Dr. Naomi Forrester, Assistant Professor, University of Texas Medical Branch, Galveston, Texas.
3. Dr. Ann Powers, Chief, Alphavirus Laboratory, Division of Vector Borne Infectious Diseases, Centers for Disease Control and Prevention, Colorado.
4. Dr. Thomas Unnasch, Professor, University of South Florida, Florida.
5. Dr. Christy Ottendorfer, Postdoctoral Fellow, Centers for Disease Control and Prevention, Colorado, Atlanta, Georgia.
6. Dr. John-Paul Mutebi, Division of Vector-Borne Diseases, Centers for Disease Control and Prevention, Colorado.
7. Dr. Marc Fischer, Professor, Division of Vector-Borne Diseases, Centers for Disease Control and Prevention, Colorado.
8. Dr. Patricia Aguilar, Assistant Professor, University of Texas Medical Branch, Galveston, Texas.
9. Dr. Aaron Brault, Chief, Flavivirus laboratory, Division of Vector Borne Infectious Diseases, Centers for Disease Control and Prevention, Colorado.
10. Dr. Philip Armstrong, Associate Scientist, Center for Vector Biology & Zoonotic Diseases, The Connecticut Agricultural Experiment Station, New Haven, Connecticut.
11. Dr. Abelardo Moncayo, Director, Vector-Borne Diseases Section, Tennessee Department of Health, Nashville, Tennessee.

Project Collaborators include.

Dr. Scott Weaver is an expert on alphavirus ecology, pathogenesis, epidemiology and evolution. Dr. Weaver will provide majority of the EEEV strains to be sequenced from his extensive collection. Dr. Weaver will also contribute to elucidating the rates of evolution, population dynamics, and the patterns of diversification and phylogeographic dispersal. These studies will be used to formulate hypotheses that can be tested both experimentally using reverse genetics, and in field epidemiology and ecology projects in Latin America.

Dr. Albert Auguste is an expert in molecular genetics and arbovirus evolution. Dr

Auguste's research interests include understanding the ecological and evolutionary factors involved in emergence, dispersal, and maintenance of zoonotic and vector-borne RNA viruses. Dr. Auguste will prepare high quality nucleic acid samples for shipping to JCVI for sequencing. As demonstrated previously, he will use phylogenetic methods to elucidate potential determinants of disease, identify adaptive mutations, as well as also infer the evolutionary dynamics, demographic and phylogeographic histories of EEEV. These studies will produce valuable data that will inform surveillance and control strategies and generate important data to test hypotheses about viral emergence and adaptation.

Dr. Robert Tesh is an expert in arbovirus characterization, epidemiology, ecology and surveillance. His research interests include virus discovery and characterization, arbovirus pathogenesis, as well as understanding the ecological and evolutionary factors involved in arbovirus emergence.

Dr. Tesh is also Director of the World Reference Center for Emerging Viruses and Arboviruses (WRCEVA) at UTMB. The WRCEVA is a direct outgrowth of the worldwide network of laboratories, established by the Rockefeller Foundation, to study the role of arthropod-borne viruses in producing human and animal disease and the mechanisms by which these viruses are maintained and transmitted in nature. When this program was initiated at the Rockefeller Foundation Virus Laboratories in New York City in 1951, fewer than 28 arboviruses had been described; and only a few, such as yellow fever, the encephalitides, and dengue were known to cause serious disease in human beings. Concurrent with the initiation of the Rockefeller Foundation program, the U.S. Army, Navy, Public Health Service and several foreign governments also established arbovirus laboratories and field research programs. This network of field laboratories relied on the Foundation's central virus reference laboratory in New York, until 1964 when the central laboratory and some of the Rockefeller staff were moved to Yale University in New Haven, and the Yale Arbovirus Research Unit (YARU) was established. In mid-1995, Drs. Robert Tesh and Robert Shope moved from YARU to UTMB in Galveston and brought the reference collection with them. The decision to establish the Reference Center at UTMB was based in part on the willingness of the University of Texas to provide modern state-of-the-art laboratory equipment and space, including BSL-3 and BSL-4 containment facilities for working with these potentially hazardous agents. A collection of more than 640 characterized type viruses is maintained with complementary sera and diagnostic antigens.

The WRCEVA at UTMB provides prompt analysis of disease outbreaks as well as identification of new and emerging viruses to agencies around the world; it also serves the world research community with basic certification of arboviruses and arboviral reagents. The extensive arbovirus reference collection maintained at the Center differs from culture collections like that of the American Type Culture Collection in several ways. First, the collection is not static; new virus strains are continually being archived. Second, instead of having just one or two prototype or well characterized strains of each known virus, there are many different strains of arboviruses of medical and veterinary importance (i.e. EEEV). A number of viruses in the collection are select agents. An effort has been made to collect representative strains of these viruses from a variety of sources, geographic localities and time periods. This diversity has proven extremely useful in studies of arbovirus emergence, evolution, pathogenesis, and now biodefense. Third, viruses and serologic reagents are provided at no cost upon request to qualified investigators; this is essential for maintaining an active reference collection, since collaborators often send us new viruses in return for this service.

Additionally, input on strain selection will be sought from the other EEEV investigators listed in section 6 above, to ensure that the genetic data produced will be relevant and appropriate for all future EEEV studies. We will also actively interact with the team at the JCVI for genome annotation and other aspects of data analysis.

6. Availability & Information of Strains:

All 100 EEEV strains to be sequenced have been retrospectively collected, stored lyophilized and are readily available for further work. Upon approval of this project, virus will be cultured, enriched for virus particles and nucleic acids prepared according to the Genome Sequencing Center's instructions. Nucleic acid samples will be prepared upon approval of the project. Attached is a list of 100 potential EEEV strains and their associated details. This list may be changed, as additional material will be sought from interested investigators listed in section 6 above. Also, selected unpassaged strains will be deep-sequenced to examine population-level quasispecies structure.

Supporting metadata include:

We will also provide relevant metadata for the EEEV strains to be studied. These include:

- Identifier
- Genus
- Species
- Specimen / Strain name
- Material type (DNA/RNA/Strain)
- Isolation source
- Date of isolation
- Location of isolation
- Organism isolated from
- Passage history (if applicable)
- Select agent status
- Other public repository location
- Other public repository identifier
- Sample provider's name for this GSC project