



SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, ML, MR, NE, SN, TD, TG).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,

Declarations under Rule 4.17:

— *of inventorship (Rule 4.17(iv))*

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

ISOLATED LYSSAVIRUS NUCLEIC ACID AND PROTEIN SEQUENCES

GOVERNMENT SPONSORSHIP

[0001] This invention was made by the Centers for Disease Control and Prevention, an agency of the United States Government.

5 CROSS REFERENCE TO RELATED APPLICATIONS

[0002] This application claims priority to U.S. Provisional Application No. 61/294,988 filed January 14, 2010, the entire contents of which are incorporated herein by reference.

FIELD OF THE INVENTION

10 [0003] The present invention relates to nucleic acid sequences of a virus. More particularly, the invention relates to a new representative species of the *Lyssavirus* genus. The invention relates to isolated nucleic acid sequences of a newly discovered Shimoni bat virus.

BACKGROUND OF THE INVENTION

15 [0004] Rabies is an ancient disease with the earliest reports possibly dated to the Old World before 2300 B.C. Rabies remains a world health threat due to remaining lack of effective control measures in animal reservoir populations and a widespread lack of human access to vaccination. Today, more than 50,000 people annually die of rabies, particularly in Asia and Africa.

[0005] Lyssaviruses are RNA viruses with single-strand, negative-sense genomes responsible for rabies-like diseases in mammals. The *Lyssavirus* genus (family *Rhabdoviridae*) includes eleven recognized species. The type species *Rabies virus* (RABV) is distributed worldwide among mammalian reservoirs - carnivores and bats. *Lagos bat virus* (LBV) circulates among pteropid bats in sub-Saharan Africa, with infrequent spill-overs into other mammals. *Mokola virus* (MOKV) has been isolated in sub-Saharan Africa from shrews, domestic cats and dogs, a rodent and two humans. The reservoir hosts of MOKV have not yet been established. *Duvenhage virus* (DUVV) was isolated from insectivorous bats and humans that died after bat bites in sub-Saharan Africa. *European bat lyssavirus, type 1* (EBLV-1) has been isolated across Europe from insectivorous bats, and the primary host species of this virus appears to be the Serotine bat (*Eptesicus serotinus*). Human cases of EBLV-1 infection have been described as well. *European bat lyssavirus, type 2* (EBLV-2) was isolated primarily from insectivorous bats of the *Myotis* genus and from humans that died after bat bites in north-western Europe. *Australian bat lyssavirus* (ABLV) circulates in Australia among insectivorous and pteropid bats,

20
25
30

and has caused at least two documented cases of human rabies. Recently, four other lyssavirus species were ratified by the International Committee on Virus Taxonomy (ICTV Official Taxonomy: Updates since the 8th Report): *Aravan virus* (ARAV) and *Khujand virus* (KHUV), isolated from insectivorous bats of the *Myotis* genus in the Central Asia; *Irkut virus* (IRKV),
5 isolated from an insectivorous bat, *Murina leucogaster* in eastern Siberia; and *West Caucasian bat virus* (WCBV), isolated from an insectivorous bat, *Miniopterus schreibersi* in south-eastern Europe. Seroprevalence to WCBV was also detected in *Miniopterus* spp. bats from Kenya, suggesting a wider geographical range than was previously believed.

[0006] Based on genetic distances, serologic cross-reactivity, and peripheral pathogenicity
10 in a mouse model, the *Lyssavirus* genus is subdivided into two phylogroups. Phylogroup I includes RABV, DUVV, EBLV-1, EBLV-2, ABLV, ARAV, KHUV and IRKV. Phylogroup II includes LBV and MOKV. The WCBV cannot be included in any of these phylogroups, and should be considered as a member of an independent phylogroup III.

[0007] The operational term 'genotype' has been used for lyssavirus classification since the
15 time when molecular techniques replaced serotyping for classification purposes. Demarcation of genotypes has been based largely on genetic distances (identity values) between members of the genus, and on the bootstrap support of phylogenetic constructions. In addition, based on identity values LBV was suggested to be sub-divided into at least two separate genotypes. However, the ICTV does not operate with viral genotypes but recognizes only viral species. Definition of a
20 viral species is complex, and cannot be based solely on genetic distances in the absence of other demarcation characteristics.

[0008] As lyssavirus surveillance in the developing countries of Africa and Asia is very limited, enhanced studies are needed to identify new lyssavirus species may be present there. This need is particularly significant because commercially available biologics do not protect
25 against lyssaviruses that do not belong to the phylogroup I. Therefore, there is the need to develop better biologics for diagnostic, treatment and prophylaxis of the diseases, caused by these viruses.

SUMMARY OF THE INVENTION

[0009] The following summary of the invention is provided to facilitate an understanding of
30 some of the innovative features unique to the present invention and is not intended to be a full description. A full appreciation of the various aspects of the invention can be gained by taking the entire specification, claims, drawings, and abstract as a whole.

[0010] The invention includes nucleic acid and protein sequences related to a newly discovered lyssavirus, the Shimoni bat virus (SHIBV). The cRNA sequence of the entire genome of SHIBV is presented herein as SEQ ID NO: 1. As such, the invention includes an isolated nucleic acid molecule including the nucleotide sequence of SEQ ID NO: 1, or a complement thereof.

[0011] The SHIBV sequence encodes five proteins. The invention also includes a portion of the nucleic acid molecule of SEQ ID NO: 1 corresponding to at least one of the lyssavirus structural genes N, P, M, G, or L. The proteins corresponding to the lyssavirus structural genes are also provided in isolated form as an isolated protein with the sequence of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, a fragment thereof, or derivative thereof.

[0012] Also provided is a process for detecting the presence of a lyssavirus virus or a nucleic acid molecule derived from a lyssavirus in a biological sample including contacting a sample with an agent that selectively binds to the a nucleic acid molecule with a sequence including at least a portion of SEQ ID NO: 1; or a protein sequence encoded by at least a portion of SEQ ID NO: 1, and detecting whether the agent binds to the virus or the nucleic acid molecule derived therefrom in the sample. The agent is optionally an antibody. In some embodiments, and agent is a nucleic acid sequence, illustratively, a nucleic acid molecule including a nucleotide sequence having between 4 and 6600 contiguous nucleotides of the nucleotide sequence of SEQ ID NO: 1, or a complement thereof.

[0013] A vaccine composition is also provided for inducing an immunological response against a lyssavirus in a subject, where the vaccine includes a pharmaceutically acceptable carrier admixed with: an isolated lyssavirus strain of SHIBV; an isolated protein sequence encoded by at least a portion of SEQ ID NO: 1, or derivative thereof; or an isolated nucleic acid sequence that has the sequence of least a portion of SEQ ID NO: 1, or derivative thereof. A vaccine is optionally formulated for parenteral or oral administration to a subject. A vaccine optionally includes an adjuvant.

[0014] Also provided is an isolated lyssavirus virus including a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1 wherein each thymine is replaced with uracil.

BRIEF DESCRIPTION OF THE DRAWINGS

[0015] FIG. 1 are phylogenetic trees of concatenated N+P+M+G+L gene sequences of lyssaviruses, obtained by the NJ method (p-distances) for nucleotide (A) and amino acid (B)

sequences, with midpoint rooting of the trees, and an unrooted ML tree for nucleotide sequences (C) where LBV lineages are indicated following Markotter et al. (2008) and Bootstrap values (1000 replicates for NJ and 100 replicates for ML) are shown for key nodes.

DETAILED DESCRIPTION OF EMBODIMENTS OF THE INVENTION

5 **[0016]** The following description of particular embodiment(s) is merely exemplary in nature and is in no way intended to limit the scope of the invention, its application, or uses, which may vary. The invention is described with relation to the non-limiting definitions and terminology included herein. These definitions and terminology are not designed to function as a limitation on the scope or practice of the invention but are presented for illustrative and descriptive
10 purposes only.

[0017] Scientific and technical terms used herein are intended to have the meanings commonly understood by those of ordinary skill in the art. Such terms are found defined and used in context in various standard references illustratively including J. Sambrook and D.W. Russell, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press; 3rd
15 Ed., 2001; F.M. Ausubel, Ed., *Short Protocols in Molecular Biology*, Current Protocols; 5th Ed., 2002; B. Alberts et al., *Molecular Biology of the Cell*, 4th Ed., Garland, 2002; D.L. Nelson and M.M. Cox, *Lehninger Principles of Biochemistry*, 4th Ed., W.H. Freeman & Company, 2004; Wild, D., *The Immunoassay Handbook*, 3rd Ed., Elsevier Science, 2005; Gosling, J. P., *Immunoassays: A Practical Approach*, Practical Approach Series, Oxford University Press,
20 2005; *Antibody Engineering*, Kontermann, R. and Dtibel, S. (Eds.), Springer, 2001; Harlow, E. and Lane, D., *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, 1988; Ausubel, F. et al., (Eds.), *Short Protocols in Molecular Biology*, Wiley, 2002; J. D. Pound (Ed.) *Immunochemical Protocols*, Methods in Molecular Biology, Humana Press; 2nd ed., 1998; B.K.C. Lo (Ed.), *Antibody Engineering: Methods and Protocols*, Methods in Molecular Biology,
25 Humana Press, 2003; and Kohler, G. and Milstein, C , *Nature*, 256:495-497 (1975); the contents of each of which are incorporated herein by reference.

[0018] The present invention includes an isolated new, previously unrecognized lyssavirus, which should be considered a new species of the *Lyssavirus* genus. Novel lyssavirus strains of the present invention are illustratively identified as Shimoni bat virus (SHIBV), fragments
30 thereof, or derivatives thereof.

[0019] The invention has utility for the development of new diagnostics, therapeutics and prophylactic therapies for viral infection. The invention also provides protein and polypeptide

sequences as well as nucleic acid sequences suitable for use as a vaccine. As such, the invention also has utility as a vaccine for the prevention of *Lyssavirus* infection.

[0020] The SHIBV was discovered in 2009 during a survey of bats in Kenya for the purpose of detecting new and potential human pathogens that could emerge from bat reservoirs within the region (Kuzmin, I.V., et al., *Emerging Infect. Dis.*, 2008; 14(12):1887-1889; and Kuzmin, I.V., et al., *J. Clin. Microbiol.*, 2008; 46(4):1451-1461). In total, 616 bats of 22 species were collected from 10 locations across Kenya including 40 sick and dead bats. The sampling and collection protocol were approved by the National Museums of Kenya and the Kenya Wildlife Service. The brain and pooled organs were collected into sterile plastic tubes. Oral swabs were placed into tubes containing Minimum Essential Medium (MEM-10, Invitrogen, Grand Island, NY). Sera were separated from blood clots by centrifugation. All samples were transported on dry ice and stored at -80°C until use.

[0021] Bat brains were subjected to the direct fluorescent antibody (DFA) test for detection of lyssavirus antigens (Dean et al., In: Meslin, F.-X., Kaplan, M.M., Koprowski, H. (Eds.). Laboratory techniques in rabies, 4th ed. WHO, Geneva, Switzerland, pp. 88-93) using monoclonal (Fujirebio Diagnostics Inc, Malvern, PA) and polyclonal (Chemicon Int., Temecula, CA) fluorescein isothiocyanate-labeled anti-rabies antibodies. When a positive DFA result was documented, the brain specimen was homogenized into 10% suspension in MEM-10, and inoculated intracranially into suckling mice (Koprowski, H., 1996. The mouse inoculation test. In: Meslin, F.-X., Kaplan, M.M., Koprowski, H. (Eds.). Laboratory techniques in rabies, 4th ed. WHO, Geneva, Switzerland, pp. 80-86) and mouse neuroblastoma (MNA) cell culture (Webster W.A., Casey, G.A., 1996. Virus isolation in neuroblastoma cell culture. In: Meslin, F.-X., Kaplan, M.M., Koprowski, H. (Eds.). Laboratory techniques in rabies, 4th ed. WHO, Geneva, Switzerland, pp. 96-104) for virus isolation. Virus sample is subjected to RT-PCR and directly sequenced.

[0022] The determined genome sequence of the Shimoni bat virus and derivatives thereof are presented. The cRNA sequence is found illustrated as SEQ ID NO: 1 as follows.

[0023] ACGCTTAACAGCAAAGTCAGAGAAGAGATAAGCCTCTACAATGAGCGAA
TCAAAATGTAACACCCCTACAATGGACTCTGAAAAGATTGTTTTCAAAGTTCGTAAC
CAGGTGGTGTCTTTGAAGCCAGAAATAATCTCTGATCAGTATGAGTACAAATATCCT
GCAATTCTGGATGGAAGAAAACCGGGGATCACTCTAGGGAGAGCTCCAGACTTAAA
CACGGCTTACAAATCAATTTTGTCTGGTATGAATGCTGCCAAATTGGATCCAGATGA
TGTGTGCTCCTATTTGGCTGCTGCGATGCAGCTGTTTGAAGGTGTTTGCCCAGAGGA

TTGGACTAGTTATGGTATTGTGATTGCAAAGAAGGGTGATAAGATCACTCCGGAAG
ATTTGATAGATGTTACTAGGACAAATGTAGAGGGAAATTGGGCTCAAACAGGGGGA
ACAGACATGACAAGAGATCCAACAAGTGCAGAACATGCATCTTTAGTGGGGTTGCT
GTTGTGCTTATATCGTCTAAGCAAGATAGTCGGCCAGAATACTGCAAATTACAAGAC
5 CAATGTGGCCGATAGAATGGAACAAATTTTTGAAACTGCACCTTTTGTGAAAATTGT
TGAACATCATACTCTAATGACTACTCATAAAATGTGTGCTAATTGGAGTACAATTCC
AAACTTTAGATTTCTTGCTGGAGCTTATGATATGTATTTTGCAAGAATAGAGCATCTT
TACTCTGCTATAAGGGTTGGAAGTGTAGTGACTGCTTACGAAGATTGTTCAAGATTG
GTCTCCTTCACTGGATTTATAAAACAAATTAATTTATCTGCTAGAGATGCTTTGCTTT
10 ATTTCTTTCACAAAAATTTTGAAGAAGAAATAAGAAGAATGTTTGAACCTGGTCAA
GAAACTGCAATTCCTCATTCTTACTTCATACATTTTAGAGCTTTGGGATTAAGCGGC
AAATCACCTTATTCATCAAATGCAGTAGGTCATACCTTCAACTTGATTCACTTTATA
GGTTGTTATATGGGTCAGATAAGATCTCTAAATGCCACAGTTATTCAAGCTTGTGCA
CCTCATGAGATGTCTGTCTTGGGAGGTTATCTAGGAGAGGAATTCTTTGGAAAAGGT
15 ACTTTTGAAAGAAGGTTCTTCAGGGATGAAAAGGAGATGCAAGACTATGCAGAATT
GGAGGGAATTAAAATAGAGGCAGCATTGGCAGATGATGGAACAGTTGACTCTGATG
ATGAGGATTTCTTTTCAGGAGAGACAAGAAGTCCAGAAGCAGTGTATAGTCGAATC
ATGATGAACAACGGCAGATTGAAAAGGTCACACATAAGGCGGTACATATCAGTCAG
CTCAAATCATCAGGCAAGGCCTAACTCATTTGCAGAGTTTTTAAACAAGGTGTACTC
20 AGATGGGTCCTAATATTGATTTCAACAGAATGCATAGGAGACAGGAGAGTTTCCATT
GTAAAAGGGGTACGAAAAAACTCAACACCCCTCTTCAATCCTTCCTCCAAGATG
AGCAAGGGGCTTATTCATCCGAGTTCCATTGCTCCGGGCTTGTGGATTTGGAGATG
GCGGAAGAACTGTCGATCTAATTCACAAAACTTGACAGACGGGCAGGCTCACTT
ACAAGGTGAACCATTAGATGTCGATTCTCTTCTGAAGATATGAGAAGAATGAGGC
25 TGATTGACATGTCCAGACAAAAGGACATCAGGATCGGAGATGAAGGAGAGAGCAG
CTCTGAGGACGAGTTTTACCTACCTAGTGGGAAAGATCCGATGGTTCCTCTTCAAGA
TTTTCTTGATGAAATCGGAGCTCAAGTAGTGAAGAGGATGAAATCTGGAGATGGGT
TCTTCAAATCTGGTCTACTGTAAGTGAAGATATAAAGGGATATATTTCTTCAAATT
TCACTGCAGCAGAACCAAGATCCAGTGACAATAAATCAGTCCAAACAGAACCTGCT
30 CAGATTCAAAAAGGGTTGCCAGAACCTCCATCCCATGAAGAGAAAGCAAAGGAAAC
ACAAGAATCCTCTAAGAACAGGCAGGAGTCAAAGCCTGCTCCCTCTTCTGATTGGG
ACAATAATCAAGAAGAAGTTGACGACATAGAGGGGGAAGTTGCTCATCAAGTTGCA
GAGAGTTTCTCTAAAAAGTACAAGTTTCCATCTAGATCATCTGGAATCTTTTTATGG

AACTTTGAACAACCTCAAGATGAATCTAGATGACATAGTGAAGGCTGCCTTGAACAT
ACCAGGAATTGATAAAATTGCAGAGAAAGGAGGGAAGCTACCATTGAGATGTATCT
TGGGATATGTTTCATTAAGTGCATCTAAGAGATTCTGCTTGCTTGCTGACAATGACA
AACTGCTAGGTTGATGCAAGATGACATCAACAATTATATTGCAAAAATTGAGGAA
5 ATAGACAAAAGTTAAAACCCCTACACAGATTCCATTGATCCACACTCAGCCTCATCAT
ATTCTtGAAAAAAACATATAACACCCCTAAAAGAATGAATTTCTGAGGAGAATTGt
GAAAAACTGCAAAGACGAAGATGCCCCAAAACCTGGGAACCTCCCTCAGCTCCTCCGG
ATGATGATGATTTGTGGCTACCTCCCCCTGAATACATGCCTTTAGCACAAATCAAAG
GTAAAGAAAGTGTGAGAACTTTTGCATAAACGGAGAGGTCAAAATCTGCAGTCCT
10 AATGGATATTCATTGAGAATAATTAGACACATCTTAAAGTCTTTTGATAACGTGTAT
TCTGGGAACAGAAGAATGATTGGACTAGCTAAAGTGGTAATCGGATTGGCACTGTC
CGGCTCACCCGTTCCAGAAGGAATGAACTGGGTTTACAAGCTAAGGAGAACACTGA
TTTTTCAATGGGCAGAATCTCAAGGCCCGTTAGAAGGGGAAGAGCTAGAATATTCC
CAGGAGATAACCTGGGACGATGAATCGGAGTTTGTGTTGTTTACAAATCAGAATCAG
15 TGCCAAACAATGTCACATCCAGGGGAGATTGTGGTGTATCAACATGAACTCCAGAG
CTTGCCAACTGTGGGCAGACATGGCCCTCAAGACCCAGCAGTCTAAAGATGATGAG
AATACTTCGCTTCTGTTGGAATGAACTTACGCATGTTAGAACTGCTTGAACTTTTAC
CTCCATTTTGAGCCGATATTATCTGCGAGCAATAGAACAAACACAGTTATGGTGCCT
GCTTTCAGCAAAATGGACACGACTTCCACCTTTATTAGTCTTTAGATTAAACTTTG
20 AGCAAGACGTGAAAAAAACATTCCCTAACGGGCTAACAGCCCTCCCCTTTTGTATC
ATGAGCAATTTATGCACCATTTTCATCCTGTGTGCAAGTATAATGGTTTCTTTAGGTG
ATTTTCCACTGTACACAATTCCCGAGAAAATAGGACCATGGACGCCCATTGATTTGA
CCCATCTGAGTTGTCCAAACAACCTGCTTTCAGAAGATGACGGATGCAGCTCTTCCT
CAACTTTCAGTTATATAGAATTGAGAACTGGTTATCTCACCCATCAAAAAGTGTCTG
25 GATTTACGTGTACTGGAGTCATCAACGAAGCAGTTACATACACCAATTTTGTAGGAT
ATGTAACAACCACTTTTAAAAGAAAACATTTCAAACCAACTGCTTCTGCCTGTAGGG
ATGCATATCATTGGAAAATTTCTGGAGATCCTCGTTATGAAGAATCACTACACACTC
CTTACCCGGACAACAGTTGGTTGAGAACAGTCACTACAACCAAAGAATCACTTCTA
ATCATCTCTCCAAGCATAGTGGAGATGGACATTTATAGTCGTTCTCTTCATTACCA
30 ATGTTTCCCACAGGGAGGTGTTATGACTTCTATAAATCTACCCCTTCTTGCTTGACCA
ACCATGACTATACAATATGGCTTCCTGATGATGCAAATGTACGACTCACATGCGATA
TCTTTGTAACGAGTACCGGAAAGAAATCAATGAATGGTTCAAAAATGTGTGGCTTTA
CTGATGAAAGAGGACTTTATAGAACTCTGAAAGGGGCTTGTAAGTTGACATTATGTG

GAAAGCCAGGACTGAGATTGTTTCGATGGCACCTGGATTTCTATCACAAGACCTGAG
ATTGTTATGTGGTGTTCCTAATCAACTTGTGAATGTGCATAACAACAGAGTGGAT
GAGATAGAACACTTAATTGTGGGTGACTTGATCCGTAGAAGAGAAGAGTGCCTGGA
TACCTTAGAGACAGTTCTTATGTCAAAGTCAGTTAGTTTTAGACGACTGAGTCATTT
5 CAGAAAGCTTGTCCCGGGTTTTGGAAAGGCATATACCATCGCCAATGGGAGCCTAA
TGGAACCAATGTACACTACAAAAGAGTAGATCGTTGGGAGGAAATATTACCCTCC
AAAGGGTGCCTAAACTGAATGACAAGTGTCTTAACCCTGAGAATGGGGTGTTTTTT
AATGGGATCATTAAAGGCCCTGATGGACAGGTTTTAATTCCAGAAATGCAGTCAAG
TTTATTGAAGCAACATATGGACTTGCTGAAAGCTTCTGTATTTCTCTAAGACATCCT
10 TTGATCGATCAAACATCTATCTTCAAAAAAGATGGAGAGGCAGATGACTTTGTTGAT
GTTACATGCCAGATCCGCATAAGTCAATATCCAATATAGATTTAGGTTTACCAGAC
TGGGGTCTATATGCAATGATTGGGGGAACTGTAGTTGCATTTCTAATATTGGTGTGT
CTCCTCTGTACTTGTTGTAAAAGAAGACGAAGGAGGAATTCCAGAAAGCCTAGCTC
AGAACAAACCCCAAAAATTTTCATCTACTCCTCCTTCTGGGACTAAAGTCATATCATC
15 TTGGGAGTCTTATAAAGGCACATCTAGTGTGTAAACTAGATATTGTTTAAAAA~~A~~ACT
CTGAGTGCCATAAATACAGCTCAACAAACTGCCTCTTTATCATCATATCACTTTATTT
TGTTTTACTCGATATAACGTTATATCTTACAGTGCATTCATCTCATATAACCTCTTAA
AATGAGACAATCAATAAAGTATATCAATCTCATCTTTTTTGTGGTACACACAACCGGG
ACAGAGTACTCATGTTGCAAGGTTATATTGGACAGAAACACCGGCAATAGACTGTG
20 ACATATTTCAATTGAGAACATTATAACAAGCAATCGATTACACCCTATCGTATGATA
CAAAAGCCGCAACTGAGATTACTTATTTACCTCGAGTGATTGTCTCAGTTCTTGAA
CCTTTAGGGTGCATATACCTTAATCAAAGCAATCTTCCTGTACAATAACTCCTTCTGA
TCTAATAATATTCAGGGATTCTGTAAATAACCTGCCTAAGAATGTAATGTAGTATAA
GAGATGTTTGGTTTTTGTTCATGTCAATTCATGTGCTAACTTTGCTCACACGTCACATT
25 GTAAGGAACA~~t~~GAAAAAAACATCTTACACTCCCCAAGATTAAGCCTCAACAGCC~~c~~TC
CTGGGTTTTGTTCCAACATGATTGAGTCCTCTGAGGTATATGATGATCCTCTTGATCC
TGCAGAGCCTGAAAGTGAATGGAGTAATACCTCCATAATTCCAAACATACTGAGAA
ATTCTGATTATAATCTGAATTCCCCTCTTTTAGAAGATCATGCTGATCTTATGTTGCA
ATGGTTGTCCTCAGGAAATCGTCCTCTGAGAATGAGTACTACAGACAATATATCCAG
30 ATCTTACAAAGTATTAAAATTGTCCTTTAAGAAAATAGATATAGCTTCAATAAAATT
TGGAGGTCAGGGAGCTCAGGCAATGATGAATACTTGGGTTCTGTGTTCTCATGCAGA
ATCCTCTAGAAGTAGGAAATGCTTGACAGACTTATCCATCTTTTACCAAAGATCTAT
ACCTATAGAGTCTATTTTAAATTACACATTAAGTAATCGAGGACTACAAACACCGAA

AGAAGGAGTATTATCATGTTTAGGTCGTATTAGTTATGATCAGTCCTTTGGTAGATA
TTTAGGGAATCTATATTCATCTTATCTATTGTTTCATACTATGATCCTTTACATGAAT
GCATTGGATTGGGAAGAAGAAAAGACGATTATAGCTTTGTGGAAAGAGATAACATC
GATAGATGTCAAGAACGACAAGGTCACCTTCAGAGATCCCTTATGGGGAAAATTTTT
5 AGTTACAAAAGAATTTGTTTACTCTTATACTAACTGCAGCTTATTTGATAGGAATTA
CACATTAATGTTGAAAGATTTGTTTTTGTCTAGATTTAACTCCTTGTTAATTCTTATA
TCACCTCCTGAACCTAGATACTCTGATGATTTAATTTCTAATCTCTGTCAATTATATA
AATCTGGGGACAAAGTCATTTTCAGAGTGTGGGAATTCCGGATACGATGTCATCAAA
ATCTTAGAACCCTACGTTGTCAATCACCTAGTTCAAAAAGCTGAGACTTTTCAGACCA
10 TTGATTCACTCTTTAGGAGACTTTCCTGCTTTCATCAAAGATAAAACCACACAGTTG
AGAGGGACTTTTGGGCCTTGTGCATCACAATTTTTCTCGGCTCTAGACCAATTTGAC
AACATTCACGATTTGGTTTTTGTGTATGGATGTTATAGACATTGGGGCCATCCTTATA
TTGACTATCGGAAAGGTCTAACCAAGTTGTTTGATCAAGTCCATATGAAAAAGACCA
TTGACAAACATTATCAGGAATGCTTGGCTAGTGATTTAGCAAAGAAAATATTGAGAT
15 GGGGATTTGATAAATATTCAAGATGGTATTTAGACACCTCATTGTTACAAAGAACC
ACCCCTTGATTCCCTATATAACAACACAGACTTGGCCTCCTAGACATGTTGTAGATT
TGTTGGGGGATTCATGGCATTCTCTCCCAATGACACAGATCTTTGAAATACCGGAAT
CAATGGATCCTTCTGAAATATTAGATGACAAATCTCACTCATTACACAAGAACAAGAC
TTGCATCCTGGCTATCTGAAAATAGAGGAGGTCCTGTCCCCAGTGAAAAAGTCATAA
20 TCACTGCCTTATCTAAGGCTCCAGTTAATCCCAGAGAGTTCTTGAAGGCGGTTGATC
TAAATGGGCTTGCTGAAGATGACTTGATCATAGGTTTGAAACCAAAGGAGAGGGAA
CTCAAATAGAAAGGCAGGTTTTTTGCATTAATGTCATGGAATTTAAGACTTTACTTT
GTTATCACGGAGAACTTTTAGCAAACCATATTCTTCCTTTATTTGATGCTTTGACGA
TGACAGACAACTTAAACAAAGTATTCAAGAAATTAATTGATAGGGTCACTGGTCAG
25 GGACTTTGTGATTACTCGAGAGTCACATATGCTTTCCATCTTGATTATGAAAAATGG
AACAATCACCAAAGACTAGAGTCTACCAAAGATGTCTTCTCTGTATTGGATAAGGCA
TTCGGACTATCCCGTGTTTTTTCTAGAACTCATGAGTTTTTTTCAGAAGTCCTGGGTTT
ATTATTCAGACAGATCTGATTTAATCGGAATTTGGAATGATCAGATATATTGTCTTG
ACATGGCCGATGGACCAACCTGCTGGAATGGTCAAGATGGTGGTCTTGAAGGTCTC
30 AGGCAAAAGGGTTGGAGTTTAGTTAGTTTGCTTATGATAGACCGAGAATCTCAAAC
AAGAAACACAAGAACCAAAAATATTAGCACAGGGTGACAATCAAGTTTTGTGTCCCA
CTTATATGTTGTCTCTGGGTTAAATAATGAGGGCTTAATGTATGAATTAGAAAGCA
TATCAAAAAATGCAATGTCAATTTATCGCGCCATGAAGAGGGGGCTTCTAAGTTGG

GACTGATAATCAAAAAGGAGGAAACCATGTGTAGTTATGACTTTTTAATCTATGGCA
AAACTCCACTGTTTAGAGGAAACATACTTGTTCCCGAGTCTAAGAGATGGGCTAGA
GTGTCCTGCATATCAAATGATCAGATAGTAACTTGGCAAATATCATGTCCACTGTA
TCAACCAATGCGCTGACTGTAGCCCAGCATTACAATCCCTAATCAAGCCTATGAGG
5 GATTTTCTTCTAATGTCTGTTCAAGCTATTTATCATTACCTCTTATTTAGCCCCATCTT
AAAAGATCGAGTATATAAAGTATTAAACTCAAAGGGGGATGACTTTCTCTTGACCAT
GTCTAGGATAATATACTTAGACCCTTCATTGGGAGGAGTGTGAGGGATGTCTCTAGG
TAGATTTTCATATAAGACAATTCTCAGATCCTGTATCAGAGGGATTGACGTTCTGGAA
AGAGATATGGCTAAGTTCTTCTGAATCTTGGATTTCATCATCTTTGTCAAGAAGCAGG
10 AAATCCTGATTTAGGAGATAGAAGTCTTGAGAGTTTTACCCGATTGCTGGAGGATCC
CACAACTTTAAACATAAGAGGGGGTGCAAGCCCTACAATTTTGTTGAAAGAAGCTA
TTAGAAAAGCCCTTTACGATGAGGTTGATAAAGTAGAAAATTCAGAATTCAGAGAG
GCTATTATCTTATCCAAAACCCACAGGGACAACCTTTATTCTATTTTTAAAATCAATTG
AACCCTATTTCCCGATTTTAAAGCGAATTGTTTCAGTTCTTCATTCTTAGGGATACC
15 AGAATCAATCATAGGTCTCATTCAAACTCCAGGACTGTTAGAAGACAATTTAGAA
AAAGCTTATCAAGAGCGTTAGAAGAATCCTTTTTTCACTCTGAAATTCATGGAATCA
GTAGAATGACTCAATCTCCTCAAAGACTTGGGAGAGTCTGGTCATGTTCCGGCAGAA
AGAGCTGATCAGCTGCGCGAGATCTCTTGGGGAAGAAAAGTTGTAGGAACAACCTGT
CCCTCACCCATCGGAGATGCTAATGTTGGTCCCCAAATCTTCAGTCTCTTGTGGTTGT
20 AGTATTAGGGAGCTGTACAACCCTAGAATCTCAGTTTCTGTTCTCCCTTCCTTTGACA
ATTCTTTCTTCTCTAGAGGATCCCTCAGGGGGTATCTAGGGTCTTCTACATCCGTCTC
AACTCAATTGTTCCATGCATGGGAGAAGGTGACCAATGTCCATGTTGTTAAAAGAGC
TCTCTCTCTCAAGGAGTCAATCAACTGGTTTATTTCCCGAGACTCTAATTTGGCACAG
ACTCTGATAAGAAACATACTGTCGTTAACAGGGCCTGAGTTTCCGATAGAAGAGGC
25 TCCAGTGTTTAAAAGGACTGGGTCTGCCCTACATAGATTCAAATCTGCAAGGTACAG
TGAAGGTGGTTATTCAGCTGTATGTCCGAATCTCTTGTCTCATATTTCTGTAAGCACT
GATACGATGTCGGATCTCACTCAAGACGGAACCTAATTTTGACTTTATGTTTCAACCC
TTAATGTTGTATGCTCAGACATGGACATCGGAGCTTGTTTCAGAAGGATTTAAGATTG
AGAGACTCAACTTTTCATTGGCATCTGAGGTGTCAAAAATGTATAAGGCCTATTGAG
30 GAGGTTAGTTTAGAAGCTCCACAAAGCTTTGCTTTTCCTGATATCTCCACTCGAATAT
CCAGAATGGTGTCTGGGGCCGTACCTCAGTTTAGAAAACTTCCTACAATAGAACTTA
AAGCGGGAGATCTTTCCAGTCTGACAAACAATGAGAGATCTTATCATGTAGGGACT
GCCCAAGGTTTGCTATATTCTATCCTCGTAGCAATACATGATCCTGGATACAATGAC

AACTCTTTATTCCCAGTCAACATATATGGAAAGGTCTCTGCTAGAAGTTATTTGAGG
GGCCTAGCAAGAGGTGTGTTTATTGGATCTTCCATATGTTTTCTAACCCGGATGACT
AATATAAATATTAACAGACCTCTTGAGTTAATTTCTGGCGTGATCTCATACATTTTGC
TCAGACTAGATAATCACCCCTTCTCTTTATGTGATGTTAAAAGAACCTGAATTGAGGT
5 CAGAGATCTTTTCAATTCCTCAGAAAATCCCAGCAGCTTACCCAACCACCATGAAGG
AAGGTAACAGATCAGTTTTATGTTATCTTCAACAGGTACTAAGATATGAAAGAGAC
AGTATGTCCTCATCTCCAGGTAACGATCTCCTGTGGATATTCTCAGACTTCAGAAGC
ATCAAAATGACATATTTGACTTTAATAACTTTCCAGGCCACCTTTGGTTACAACGG
ATTGAGAGAAATTTGTCTAAACAAATGAGAGCAAACTTCGTCAACTGAACTCATT
10 AATGAGACAAGTTTTGGGCGGTCATGGAGAGGAAAATATCGAATCGGATGATGAAA
TAAACTCTCTGCTCAAGGAAGCTCTTAGAAGAAGTAGATGGGTAGATCAGGAGGTA
CGTCATGCTGCTAAGACTATGAGACCTGACTTAAGTCCAGTTCCAAAGAACTCACGC
AAAATTGGTTCTTCAGAATGGATTTGCTCTGCCCAACAAATAGCTTTCTCAACCTCTT
TGAACCCTGCCCCAATGTCAGATATAGATTTGAGACTATTGTCTCGACAGTTTCAAA
15 ACCCTCTTATGTCAGGTTTAAGAGTCGTCCAATGGGCTACAGGAGACATTATAAGT
TAAAGCCAATCTTGGATGATTTAGAAGTCGGTCCAAGCTTGTCACTCGTAGTAGGAG
ATGGTTCAGGGGGAATTTCTAGGACTGTTCTGAACATGTTTCCAGACTCTAAGTTGGT
TTTCAATAGCCTTCTAGAGGTAAATGATTTGATGGCTTCGGGAACACATCCTCTTCCT
CCTTCCGCCTTGATGAGAGGAGGAGAAGATATAACATCAAGAGTGATAGACTTTGA
20 ATCTATATGGGAAAAGCCTTCTGATTTAAGAAATTCTGTCACATGGAAATATTTTCA
GTCTGTCCAAACAAGAGTGAAGGCTCATTTTGATCTAATCGTTTGTGATGCAGAGGT
AACAGACATTGAATCTGTCAACAAGATCACTTTGCTTTTGTCTGATTTTGCAATGTCA
ATAAGAGGACCCTTATGTCTGATCTTCAAGACATATGGAACATATGCTGATTAATCCT
GATTACAAGGCCATCCATCACTTATCCAGAGCATTTCCTAACATGATTGGGTTTGTT
25 ACTCAGATGACATCATCTTTCTCTTCTGAAGTTTATCTCAGATTTTCCAAAAAGGGTC
ATTTTTTTAGAGAGCATGAGTCACTCACTGCCTCCACTATTAGAGAAATGAGCCTGG
TTTTGTTCAATTGTAGCAATCCTAAAAGTGAAATGTTGAGGGCCAGATCGCTGAACT
ATCAAGATTTGATCAGAGGATTCCCTGAAGAGATAATCTCCAATCCATATAACGAA
ATGATAATAACTTTAATTGATAGTGAAGTCGAGTCATTTCTTGTCATAAGTTGGTG
30 GATGATCTAGAGTTAAAAAGAGGCTCCTCATCCAAGATGGCAATTATTGTTGCAATA
ATTATACTGTTTTCTAACAGGGTCTTCAATGTGTCAAAGTCAATTAAAGATACAAAA
TTCTTTCCTCCCTCAGATCCGAAAATTCTCCGACACTTTAACATATGTTTGAGCACGA
TGTTGTTTTTGTCTACAACATATGGGGGACCTGTCAAACCTTCACAAAAATTCATGAGT

TGTATAACTCTCCAGTCATATACTATTTTCGAAAACAGACTATCAAAGGCCAAAAAGT
 TTCTCTCTTGGAGTTGGGCTGATCCCAGTAGCATATTCAAAAAGGTATCTTGTA
 CATCTCTTAGTTTATCAGCTCATTGGATTAGAATGATTTATAAAATTGTCAAACTAC
 TAGGTTGAACTCAAATCCCCGAGAGTTGTTGAAAGAGGTAGAAATTTACCTTAAGG
 5 GATACAATCGATGGATTACAATGAGAGATGTTAGATCTAGAACCTCCTTGCTAGACT
 ATTGTTGTTTGTAAATTTGAATCGTTGAGGATTCTCACCTAAAAGTTCAACTGACAGG
 TCACCTGGACTGAAGAAGATACAAAACCTAGAAAAAACATGTCTAAGACTCACAGC
 ATGTCACTGCTTGAGGACTCTTCTTTTTTATGTTGTTTTTTTGTTAAGCGT (SEQ ID
 NO: 1)

10 **[0024]** The negative-sense lyssavirus genome encodes five proteins: the nucleoprotein (N);
 phosphoprotein (P); matrix protein (M); glycoprotein (G); and RNA polymerase (L) that are
 present in the overall genome in the order 3'-N-P-M-G-L-5'. Tordo, N, et al., *Proc Natl Acad Sci*
USA, 1986; 83:3914-3918. The invention also includes isolated nucleic acids encoding to the
 protein sequences of the five structural genes (N, P, M, G and L) of the SHIBV virus and
 15 derivatives thereof. The present invention also includes isolated protein sequences of the SHIBV
 virus corresponding to the five structural genes (N, P, M, G and L).

[0025] Also provided are isolated nucleic acids encoding the structural proteins of SHIBV.
 A nucleic acid sequence encoding nucleoprotein (N gene) includes nucleotides 71 to 1423 of
 SEQ ID NO: 1, or derivatives thereof. A nucleic acid sequence encoding the phosphoprotein (P
 20 gene) includes nucleotides 1522 to 2439 of SEQ ID NO: 1, and derivatives thereof. A nucleic
 acid sequence encoding the matrix protein (M gene) includes nucleotides 2516 to 3124 of SEQ
 ID NO: 1, and derivatives thereof. A nucleic acid sequence encoding the glycoprotein (G gene)
 includes nucleotides 3330 to 4898 of SEQ ID NO: 1, and derivatives thereof. A nucleic acid
 sequence encoding the polymerase protein (L gene) includes nucleotides 5512 to 11895 of SEQ
 25 ID NO: 1, and derivatives thereof.

[0026] As used herein, a derivative of a nucleic acid is a sequence of nucleotides that is not
 found in nature. A derivative of a nucleic acid may include at least one substitution, deletion, or
 modification of at least one base. A derivative of a nucleic acid is optionally 50% or more
 identical to the wild-type sequence. Optionally a derivative is has identity relative to the wild-
 30 type sequence of 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%,
 96%, 97%, 98%, 99%, or more, or any value or range of values therebetween. Base
 modifications are modifications known in the art that may alter the melting temperature, alter
 the annealing temperature, alter the translated amino acid sequence, shorten or lengthen an

encoded amino acid sequence relative to wild-type, alter expression level of mRNA or protein, or other modification effect known in the art.

[0027] The term "nucleotide" is intended to mean a base-sugar-phosphate combination either natural or synthetic, linear, circular and sequential arrays of nucleotides and nucleosides, e.g. cDNA, genomic DNA, mRNA, and RNA, oligonucleotides, oligonucleosides, and derivatives thereof. Included in this definition are modified nucleotides which include additions to the sugar-phosphate groups as well as to the bases.

[0028] The term "nucleic acid" or "oligonucleotide" refers to multiple nucleotides attached in the form of a single or double stranded polynucleotide that can be natural, or derived synthetically, enzymatically, and by cloning methods. The terms "nucleic acid" and "oligonucleotide" may be used interchangeably in this application.

[0029] An inventive nucleic acid sequence is provided. The nucleic acid sequence relates to the genome sequence of Shimoni bat virus as found in SEQ ID NO: 1, portions thereof, and derivatives thereof. The inventive nucleic acid sequence is optionally isolated from the cellular or viral materials with which it is naturally associated. Numerous methods are known in the art for the synthesis and production of nucleic acid sequences illustratively including cloning and expression in cells such as *E. coli*, insect cells such as Sf9 cells, yeast, and mammalian cell types such as Hela cells, Chinese hamster ovary cells, or other cells systems known in the art as amendable to transfection and nucleic acid and/or protein expression. Methods of nucleic acid isolation are similarly recognized in the art. Illustratively, plasmid DNA amplified in *E. coli* is cleaved by suitable restriction enzymes such as NdeI and XhoI to linearize glycoprotein DNA. The glycoprotein DNA is subsequently isolated following gel electrophoresis illustratively using a S.N.A.P.TM UV-Free Gel Purification Kit (Invitrogen, Carlsbad, CA) as per the manufacturer's instructions.

[0030] Numerous agents are commercially available to facilitate cell transfection illustratively including synthetic or natural transfection agents such as lipid agents illustratively including LIPOFECTIN recognized in the art as a 1:1 (w/w) liposome formulation of the cationic lipid N-[1-(2,3-dioleoyloxy)propyl]-n,n,n-trimethylammonium chloride (DOTMA) and dioleoyl phosphatidylethanolamine (DOPE) in membrane filtered water, baculovirus, naked plasmid or other DNA, or other systems known in the art.

[0031] The nucleotide sequences of the invention may be isolated by conventional uses of polymerase chain reaction or cloning techniques such as those described in conventional texts. For example, the nucleic acid sequences of this invention may be obtained via reverse

transcription of viral RNA, and subsequent amplification and sequencing of the cDNA, using DNA primers and probes and PCR techniques. These sequences, fragments thereof, modifications thereto and the full-length sequences may be constructed recombinantly using conventional genetic engineering or chemical synthesis techniques or PCR, and the like.

- 5 **[0032]** The present invention also encompasses isolated proteins derived from SHIBV genetic sequence or derivatives of proteins derived from SHIBV. In some embodiments, a protein is a SHIBV nucleoprotein, phosphoprotein, matrix protein, glycoprotein, polymerase protein, or derivatives of any of these proteins. The protein is optionally recombinant. The protein is optionally isolated. However, it is also envisioned that naturally occurring structural
10 protein is optionally isolated from at least a portion of sample material from which the wild-type sequence is normally found. Methods for purification of protein from organism derived samples are known and are within the level of skill in the art.

[0033] An inventive protein illustratively has an amino acid sequence of nucleoprotein as represented by SEQ ID NO: 2, or derivatives thereof.

- 15 **[0034]** MDSEKIVFKVRNQVVSLKPEIISDQYEEKYPAILDGRKPGITLGRAPDLNTAY
KSILSGMNAAKLDPDDVCSYLAAMQLFEGVCPEDWTSYGIVIAKKGDKITPEDLIDVT
RTNVEGNWAQTGGTDMTRDPTSAEHASLVGLLLCLYRLSKIVGQNTANYKTNVADRM
EQIFETAPFVKIVEHHTLMTTHKMCANWSTIPNFRFLAGAYDMYFARIEHLYSAIRVGTV
VTAYEDCSGLVSFTGFIKQINLSARDALLYFFHKNFEEEIRRMFEPGQETAIPHSYFIHFRA
20 LGLSGKSPYSSNAVGHTFNLIHFIGCYMGQIRSLNATVIQACAPHEMSVLGGYLGEFFG
KGTFFERRFFRDEKEMQDYAELEGIKIEAALADDGTVDSDDDEFFSGETRSPPEAVYSRIM
MNNGLRLKRSHIRRYISVSSNHQARPNSFAEFLNKVYSDGS (SEQ ID NO: 2)

[0035] An inventive protein illustratively has an amino acid sequence of phosphoprotein as represented by SEQ ID NO: 3, or derivatives thereof.

- 25 **[0036]** MSKGLIHPSSIRSLVDLEMAEETVDLIHKNLTDGQAHLQGEPLDVDLDPED
MRRMRLIDMSRQKDIRIGDEGESSSEDEFYLP SGKDPMVPLQDFLDEIGAQVVKRMKSG
DGFFKIWSTVTEDIKGYISSNFTAAEPRSSDNKSVQTEPAQIQKGLPEPPSHEEKAKETQE
SSKNRQESKPAPSSDWDNNQEE VDDIEGE VAHQVAESFSKKYKFPSRS SGIFLWNFEQL
KMNLDIVKAALNIPGIDKIAEKGGKLPLRCILGYVSLSASKRFCLLADNDKTARLMQD
30 DINNYIAKIEEIDKS (SEQ ID NO: 3)

[0037] An inventive protein illustratively has an amino acid sequence of matrix protein represented by SEQ ID NO: 4, or derivatives thereof.

[0038] MNFLRRIVKNCKDEDAPKLGTPSAPPDDDDLWLPPPEYMPLAQIKGKESVR
NFCINGEVKICSPNGYSFRIIRHILKSFDNVYSGNRRMIGLAKVVIGLALSGSPVPEGMNW
VYKLRRTLIFQWAESQGPLEGEELEYSQEITWDDSEFVGLQIRISAKQCHIQRLLWCIN
MNSRACQLWADMALKTQQSKDDENTSLLLE (SEQ ID NO: 4)

- 5 **[0039]** An inventive protein illustratively has an amino acid sequence of glycoprotein represented by SEQ ID NO: 5, or derivatives thereof.

[0040] MSNLCTIFILCASIMVSLGDFPLYTIPEKIGPWTPIDLTHLSCPNNLLSEDDGCS
SSSTFSYIELRTGYLTHQKVSGFTCTGVINEAVTYTNFVGYYVTTTFKRKHFKPTASACRD
AYHWKISGDPRYEESLHTPYPDNSWLRTVTTTTKESLLIISPSIVEMDIYSRSLHSPMFPTG
10 RCYDFYKSTPSCLTNHDYTIWLPDDANVRLTCDIFVTSTGKKSMNGSKMCGFTDERGL
YRTLKGACKLTLCGKPGRLRLFDGTWISITRPEIVMWCSNPQLVNVHNNRVDEIEHLIVG
DLIRREECLDTLETVLMKSVSFRRLSHFRKLVPGFGKAYTIANGSLMETNVHYKRVD
RWEEILPSKGCLKLNDKCLNPENG VFFNGIIGKPGDGQVLIPEMQSSLLKQHMDLLKASVF
PLRHPLIDQTSIFKKDGEADDFVDVHMPDPHKSISNIDLGLPDWGLYAMIGGTVVAFLLIL
15 VCLLCTCCKRRRRRRNSRKPSSEQTPKISSTPPSGTKVISSWESYKGTSSV (SEQ ID NO: 5)

- [0041]** An inventive protein illustratively has an amino acid sequence of polymerase protein represented by SEQ ID NO: 6, or derivatives thereof.

[0042] MIESSEVYDDPLDPAEPESEWSNTSIIPNILRNSDYNLNSPLLEDHADLMLQW
LSSGNRPLRMSTTDNISRSYKVLKLSFKKIDIASIKFGGQGAQAMMNTWVLCSHAESSRS
20 RKCLTDLSIFYQRSIPIESILNYTLNRLGLQTPKEGVLSCGRISYDQSFGRYLGNLYSSYL
LFHTMILYMNALDWEEETIHALWKEITSIDVKNDKVTFRDPLWGKFLVTKEFVYSYTN
CSLFDRNYTLMKDLFLSRFNSLLILISPEPRYSDDLISNLCQLYKSGDKVISECGNSGY
DVIKILEPYVVNHLVQKAETFRPLIHSLGDFPAFIKDKTTQLRGTFGPCASQFFSALDQFD
NIHDLVFVYGCYRHWGHPYIDYRKGLTKLFDQVHMKKTIDKHYYQECLASDLAKKILR
25 WGFDKYSRWYLDTSLLPKNHPLIPYITTQTWPPRHVVDLLGDSWHSPLMTQIFEIPESM
DPSEILDDKSHSFTRTRLASWLSENRGGPVPSEKVIITALSKAPVNPREFLKAVDLNLGLAE
DDLIIGLKPKERELKIEGRFFALMSWNLRLYFVITEKLLANHILPLFDALTMTDNLNKFV
KKLIDRVTGQGLCDYSRVTYAFHLDYEKWNHQRLESTKDVFSVLDKAFGLSRVFSRT
HEFFQKSWVYYSDRSDLIGIWNQDIYCLDMADGPTCWNGQDGGLEGLRQKGWSLVSL
30 LMIDRESQTRNTRTKILAQGDNQVLCPTYMLSSGLNNEGLMYELESISKNAMSIYRAIEE
GASKLGLIIKKEETMCSYDFLIYGKTPLFRGNILVPESKRWARVSCISNDQIVNLANIMST
VSTNALTVAQHSQSLIKPMRDFLLMSVQAIYHYLLFSPILKDRVYKVLNSKGDDFLLTM
SRIIYLDPSLGGVSGMSLGRFHIRQFSDPVSEGLTFWKEIWLSSSESWIHHLCCQEAGNPDL

GDRSLESFTRLLEDPTTLNIRGGASPTILLKEAIRKALYDEVKVENSEFREAIILSKTHRD
 NFILFLKSIEPLFPRFLSELFSSSFLGIPESIIGLIQNSRTVRRQFRKSLSRALEESFFHSEIHGI
 SRMTQSPQRLGRVWSCSAERADQLREISWGRKVVGTTVPHHPSEMLMLVPKSSVSCGCSI
 RELYNPRISVSVLPSFDNSFFSRGSLRGLGSSTSVSTQLFHAWEKVTNVHVVKRALSLK
 5 ESINWFISRDSNLAQTLIRNILSLTGPEFPIEEAPVFKRTGSALHRFKSARYSEGGYSAVCP
 NLLSHISVSTDMSDLTQDGTNFDNFMFQPLMLYAQTTWSELVQKDLRLRDSTFWHLR
 CQKCIPIEEVSLEAPQSFAPDISTRISRMVSGAVPQFRKLPTIELKAGDLSSLTNNERSY
 HVGTAQGLLYSILVAIHDPGYNDNSLFPVNIYGKVSARSYLRLARGVFIGSSICFLTRM
 TNININRPLELISGVISYILLRLDNHPSLYVMLKEPELRSEIFSIPQKIPAAAYPTTMKEGNRS
 10 VLCYLQQVRLRYERDSMSSSPGNDLLWIFSDFRSIKMTYLTITFQAHLWLQRIERNLSKQ
 MRAKLRQLNSLMRQVLGGHGEENIESDDEINSLLEALRRTRWVDQEVRAAKTMRP
 DLSPVPKNSRKIGSSEWICSAQQIAFSTSLNPAPMSDIDLRLLSRQFQNPLMSGRLRVVQW
 ATGAHYKLKPILDDLEVGPSLSLVVGDGSGGISRTVLNMFPSKLVFNSLLEVNDLMAS
 GTHPLPPSALMRGGEDITSRVIDFESIWEKPSDLRNSVTWKYFQSVQTRVKAHFDLIVCD
 15 AEVTDIESVNKITLLLSDFAMSIRGPLCLIFKTYGTMLINPDYKAIHHLSRAFPNMIGFVT
 QMTSSFSSEVYLRFSKKGHFFREHESLTASTIREMSLVLFNCSNPKSEMLRARSLEYQDLI
 RGFPEEIIISNPYNEMIITLIDSEVESFLVHKLVDDELEKRGSSSKMAIIVAIILFSNRVFNVS
 KSIKDTKFFPPSDPKILRHFNICLSTMLFLSTTMGDLSNFTKIHLYNSPVIYYFRKQTIKG
 KKFLSWSWADPSSIFKKVSCNSSLSLSAHWIRMIYKIVKTTRLNSNPPELLKEVEIYLGK
 20 YNRWITMRDVRSTSLLDYCCL (SEQ ID NO: 6)

[0043] Illustrative nucleic acid and protein sequences of the present invention can be found at GenBank Accession No: GU170201, the entire contents of which are incorporated herein by reference.

[0044] The description is directed primarily to the glycoprotein (G) protein of Shimoni bat virus (SHIBV) as an illustrative example of a protein sequence that is the target of antibodies and useful for development of a vaccine to SHIBV. It is to be appreciated that the use of glycoprotein is for exemplary purposes only and not meant to be a limitation on the isolated protein structures of the present invention, which are equally appreciated to include the isolated protein sequences encoded by the genes N, P, M, G, and L. All proteins encoded by the
 25 nucleotide sequence of Shimoni bat virus as in SEQ ID NO: 1, portions thereof, or derivatives thereof, are similarly included.

[0045] The terms "polypeptide," "peptide," and "protein" are synonymous as used herein and are intended to mean a natural or synthetic compound containing two or more amino acids.

Amino acids present in a protein include the common amino acids alanine, cysteine, aspartic acid, glutamic acid, phenylalanine, glycine, histidine, isoleucine, lysine, leucine, methionine, asparagine, proline, glutamine, arginine, serine, threonine, valine, tryptophan, and tyrosine as well as less common naturally occurring amino acids, modified amino acids or synthetic compounds, such as alpha-asparagine, 2-aminobutanoic acid or 2-aminobutyric acid, 4-aminobutyric acid, 2-aminocaproic acid (2-aminodecanoic acid), 6-aminocaproic acid, alpha-glutamine, 2-aminoheptanoic acid, 6-aminohexanoic acid, alpha-aminoisobutyric acid (2-aminoalanine), 3-aminoisobutyric acid, beta-alanine, allo-hydroxylysine, allo-isoleucine, 4-amino-7-methylheptanoic acid, 4-amino-5-phenylpentanoic acid, 2-aminopimelic acid, gamma-amino-beta-hydroxybenzenepentanoic acid, 2-aminosuberic acid, 2-carboxyazetidine, beta-alanine, beta-aspartic acid, biphenylalanine, 3,6-diaminohexanoic acid, butanoic acid, cyclobutyl alanine, cyclohexylalanine, cyclohexylglycine, N5-aminocarbonylornithine, cyclopentyl alanine, cyclopropyl alanine, 3-sulfoalanine, 2,4-diaminobutanoic acid, diaminopropionic acid, 2,4-diaminobutyric acid, diphenyl alanine, N,N-dimethylglycine, diaminopimelic acid, 2,3-diaminopropanoic acid, S-ethylthiocysteine, N-ethylasparagine, N-ethylglycine, 4-azaphenylalanine, 4-fluoro-phenylalanine, gamma-glutamic acid, gamma-carboxyglutamic acid, hydroxyacetic acid, pyroglutamic acid, homoarginine, homocysteic acid, homocysteine, homohistidine, 2-hydroxyisovaleric acid, homophenylalanine, homoleucine, homoproline, homoserine, homoserine, 2-hydroxypentanoic acid, 5-hydroxylysine, 4-hydroxyproline, 2-carboxyoctahydroindole, 3-carboxyisoquinoline, isovaline, 2-hydroxypropanoic acid (lactic acid), mercaptoacetic acid, mercaptobutanoic acid, sarcosine, 4-methyl-3-hydroxyproline, mercaptopropanoic acid, norleucine, nipecotic acid, nortyrosine, norvaline, omega-amino acid, ornithine, penicillamine (3-mercaptovaline), 2-phenylglycine, 2-carboxypiperidine, sarcosine (N-methylglycine), 2-amino-3-(4-sulfophenyl)propionic acid, 1-amino-1-carboxycyclopentane, 3-thienylalanine, epsilon-N-trimethyllysine, 3-thiazolylalanine, thiazolidine 4-carboxylic acid, alpha-amino-2,4-dioxypyrimidinepropanoic acid, and 2-naphthylalanine. Accordingly, the term "protein" as used herein includes peptides having between 2 and about 2500 amino acids or having a molecular weight in the range of about 150 - 250,000 Daltons.

[0046] A protein is obtained by any of various methods known in the art illustratively including isolation from a cell or organism, chemical synthesis, expression of a nucleic acid and partial hydrolysis of proteins. Chemical methods of peptide synthesis are known in the art and include solid phase peptide synthesis and solution phase peptide synthesis for instance, or by the

method of Hackeng, TM, et al., *Proc Natl Acad Sci USA*, 1997; 94(15):7845-50, the contents of which are incorporated herein by reference.

[0047] A protein included in an inventive composition or method may be a naturally occurring or non-naturally occurring protein. The term "naturally occurring" refers to a protein endogenous to a cell, tissue or organism (in this case a lyssavirus) and includes allelic variations. A non-naturally occurring protein is synthetic or produced apart from its naturally associated organism or modified and is not found in an unmodified cell, tissue or organism.

[0048] Modifications and changes can be made in the structure of the proteins that are the subject of the application and still obtain a molecule having similar characteristics as the wild-type protein. For example, certain amino acids can be substituted for other amino acids in a sequence without appreciable loss of activity or ability to interact with other proteins, lipids, carbohydrates, cofactors, or other molecules. Because it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence and nevertheless obtain a protein with like properties. For example, one or more amino acid substitutions, additions, or deletions can be made without altering the functional properties of SHIBV proteins. It is also appreciated that several mutations optionally increase, decrease, or do not change the immunogenicity of an inventive protein.

[0049] Conservative amino acid substitutions can be made in SHBIV polypeptides to produce derivatives. Conservative amino acid substitutions are art recognized substitutions of one amino acid for another amino acid having similar characteristics. For example, each amino acid may be described as having one or more of the following characteristics: electropositive, electronegative, aliphatic, aromatic, polar, hydrophobic and hydrophilic. A conservative substitution is a substitution of one amino acid having a specified structural or functional characteristic for another amino acid having the same characteristic. Acidic amino acids include aspartate, glutamate; basic amino acids include histidine, lysine, arginine; aliphatic amino acids include isoleucine, leucine and valine; aromatic amino acids include phenylalanine, glycine, tyrosine and tryptophan; polar amino acids include aspartate, glutamate, histidine, lysine, asparagine, glutamine, arginine, serine, threonine and tyrosine; and hydrophobic amino acids include alanine, cysteine, phenylalanine, glycine, isoleucine, leucine, methionine, proline, valine and tryptophan; and conservative substitutions include substitution among amino acids within each group. Amino acids may also be described in terms of relative size, alanine, cysteine,

aspartate, glycine, asparagine, proline, threonine, serine, valine, all typically considered to be small.

[0050] In making such changes, the hydropathic index of amino acids can be considered. The importance of the hydropathic amino acid index in conferring interactive biologic function on a polypeptide is generally understood in the art. It is known that certain amino acids can be substituted for other amino acids having a similar hydropathic index or score and still result in a polypeptide with similar biological activity. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics. Those indices are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cysteine (+2.5); methionine (+1.9); alanine (+1.8); glycine (-0.4); threonine (-0.7); serine (-0.8); tryptophan (-0.9); tyrosine (-1.3); proline (-1.6); histidine (-3.2); glutamate (-3.5); glutamine (-3.5); aspartate (-3.5); asparagine (-3.5); lysine (-3.9); and arginine (-4.5).

[0051] It is believed that the relative hydropathic character of the amino acid determines the secondary structure of the resultant polypeptide, which in turn defines the interaction of the polypeptide with other molecules, such as enzymes, substrates, receptors, antibodies, antigens, and the like. It is known in the art that an amino acid can be substituted by another amino acid having a similar hydropathic index and still obtain a functionally equivalent polypeptide. In such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

[0052] Substitution of like amino acids can also be made on the basis of hydrophilicity, particularly, where the biological functional equivalent polypeptide or peptide thereby created is intended for use in immunological embodiments. The following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 $\pm$ 1); glutamate (+3.0 $\pm$ 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); proline (-0.5 $\pm$ 1); threonine (-0.4); alanine (-0.5); histidine (-0.5); cysteine (-1.0); methionine (-1.3); valine (-1.5); leucine (-1.8); isoleucine (-1.8); tyrosine (-2.3); phenylalanine (-2.5); tryptophan (-3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent polypeptide. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

[0053] As outlined above, amino acid substitutions are generally based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include (original residue: exemplary substitution): (Ala: Gly, Ser), (Arg: Lys), (Asn: Gin, His), (Asp: Glu, Cys, Ser), (Gin: Asn), (Glu: Asp), (Gly: Ala), (His: Asn, Gin), (He: Leu, Val), (Leu: lie, Val), (Lys: Arg), (Met: Leu, Tyr), (Ser: Thr), (Thr: Ser), (Tip: Tyr), (Tyr: Trp, Phe), and (Val: He, Leu). Embodiments of this disclosure thus contemplate functional or biological equivalents of a polypeptide as set forth above. In particular, embodiments of the polypeptides can include variants having about 50%, 60%, 70%, 80%, 90%, and 95% sequence identity to the polypeptide of interest.

[0054] An inventive glycoprotein is illustratively recombinant. An inventive protein may be coexpressed with associated tags, modifications, other proteins such as in a fusion peptide, or other modifications or combinations recognized in the art. Illustrative tags include 6x His, FLAG, biotin, ubiquitin, SUMO, or other tag known in the art. A tag is illustratively cleavable such as by linking to glycoprotein or an associated protein via an enzyme cleavage sequence that is cleavable by an enzyme known in the art illustratively including Factor Xa, thrombin, SUMOstar protein as obtainable from Lifesensors, Inc., Malvern, PA, or trypsin. It is further appreciated that chemical cleavage is similarly operable with an appropriate cleavable linker.

[0055] Protein expression is illustratively accomplished from transcription of nucleic acid sequence encoding glycoprotein or derivatives thereof, translation of RNA transcribed from glycoprotein nucleic acid sequence such as a cDNA sequence, modifications thereof, or fragments thereof. Protein expression is optionally performed in a cell based system such as in *E. coli*, Hela cells, or Chinese hamster ovary cells. It is appreciated that cell-free expression systems are similarly operable.

[0056] It is recognized that numerous derivatives of lyssavirus proteins are within the scope of the present invention including amino acid substitutions, alterations, modifications, or other amino acid changes that increase, decrease, or do not alter the function of the protein sequence. A derivative of a protein sequence is a sequence of amino acids that is not found in nature. Several post-translational modifications are similarly envisioned as within the scope of the present invention illustratively including incorporation of a non-naturally occurring amino acid, phosphorylation, glycosylation, addition of pendent groups such as biotin, fluorophores, lumiphores, radioactive groups, antigens, or other molecules.

[0057] The present invention also provides a vector with an inventive nucleic acid sequence therein wherein the nucleic acid sequence optionally encodes an inventive protein. Illustrative vectors include a plasmid, cosmid, cationic lipids, non-liposomal cationic vectors, cationic cyclodextrin, viruses with RNA or DNA genetic material, polyethylenimines, histidylated polylysine, or other vector system known in the art. A vector is optionally a plasmid. A suitable vector optionally possesses cell type specific expression or other regulatory sequences or sequences operable to stimulate or inhibit gene or protein expression. A vector illustratively contains a selection marker such as an antibiotic resistance gene.

[0058] Also provided is a host cell transformed with an appropriate vector or with the inventive nucleic acid sequence, optionally that encodes an inventive protein or derivative. A host cell for expression of polypeptides and fragments thereof can be prokaryotic or eukaryotic, such as bacterial, plant, insect, fungus, yeast, and mammalian cells. Optional host cells include *E. coli* or Sf9 cells. An expression vector is introduced into a host cell using well-known techniques such as infection or transfection, including calcium phosphate transfection, liposome-mediated transfection, electroporation and sonoporation.

[0059] A method is also provided for recombinantly expressing a inventive nucleic acid or protein sequence or fragments thereof wherein a cell is transformed with an inventive nucleic acid sequence and cultured under suitable conditions that permit expression of nucleic acid sequence or protein either within the cell or secreted from the cell. Cell culture conditions are particular to cell type and expression vector. Culture conditions for particular vectors and cell types are within the level of skill in the art to design and implement without undue experimentation. Techniques for the expression and purification of recombinant proteins are known in the art (see Sambrook Eds., *Molecular Cloning: A Laboratory Manual* 3rd ed. (Cold Spring Harbor, N.Y. 2001).

[0060] Some embodiments of the present invention are compositions containing nucleic acid that can be expressed as encoded polypeptides or proteins. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression. It is believed that virtually any expression system may be employed in the expression of the claimed nucleic and amino sequences.

[0061] As used herein, the terms "engineered" and "recombinant" cells are synonymous with "host" cells and are intended to refer to a cell into which an exogenous DNA or RNA segment or gene, such as a cDNA or gene has been introduced. Therefore, engineered cells are

distinguishable from naturally occurring cells which do not contain a recombinantly introduced exogenous DNA segment or gene. A host cell is optionally a naturally occurring cell that is transformed with an exogenous DNA segment or gene or a cell that is not modified.

[0062] To express a recombinant encoded polypeptide in accordance with the present invention one optionally prepares an expression vector that includes a nucleic acid under the control of one or more promoters. To bring a coding sequence "under the control of a promoter, one positions the 5' end of the translational initiation site of the reading frame generally between about 1 and 50 nucleotides "downstream" of (i.e., 3' of) the chosen promoter. The "upstream" promoter stimulates transcription of the inserted DNA and promotes expression of the encoded recombinant protein. This is the meaning of "recombinant expression" in the context used here.

[0063] Many standard techniques are available to construct expression vectors containing the appropriate nucleic acids and transcriptional/translational control sequences in order to achieve protein or peptide expression in a variety of host-expression systems. Cell types available for expression include, but are not limited to, bacteria, such as *E. coli* and *B. subtilis* transformed with recombinant phage DNA, plasmid DNA or cosmid DNA expression vectors.

[0064] Certain examples of prokaryotic hosts are *E. coli* strain RR1, *E. coli* LE392, *E. coli* B, *E. coli* .chi. 1776 (ATCC No. 31537) as well as *E. coli* W3110 (F-, lambda-, prototrophic, ATCC No. 273325); bacilli such as *Bacillus subtilis*; and other enterobacteriaceae such as *Salmonella typhimurium*, *Serratia marcescens*, and various *Pseudomonas* species.

[0065] In general, plasmid vectors containing replicon and control sequences that are derived from species compatible with the host cell are used in connection with these hosts. The vector optionally carries a replication site, as well as marking sequences that are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is often transformed using pBR322, a plasmid derived from an *E. coli* species. Plasmid pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid or phage may also contain, or be modified to contain, promoters that can be used by the microbial organism for expression of its own proteins.

[0066] In addition, phage vectors containing replicon and control sequences that are compatible with the host cell can be used as transforming vectors in connection with these hosts. For example, the phage lambda may be utilized in making a recombinant phage vector that can be used to transform host cells, such as *E. coli* LE392.

[0067] Further useful vectors include pIN vectors and pGEX vectors, for use in generating glutathione S-transferase (GST) soluble fusion proteins for later purification and separation or cleavage. Other suitable fusion proteins are those with β -galactosidase, ubiquitin, or the like.

[0068] Promoters that are most commonly used in recombinant DNA construction include the β -lactamase (penicillinase), lactose and tryptophan (trp) promoter systems. While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling those of skill in the art to ligate them functionally with plasmid vectors.

[0069] For expression in *Saccharomyces*, the plasmid YRp7, for example, is commonly used. This plasmid contains the trp1 gene, which provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example ATCC No. 44076 or PEP4-1. The presence of the trp1 lesion as a characteristic of the yeast host cell genome then provides an effective environment for detecting transformation by growth in the absence of tryptophan.

[0070] Suitable promoting sequences in yeast vectors include the promoters for 3-phosphoglycerate kinase or other glycolytic enzymes, such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. In constructing suitable expression plasmids, the termination sequences associated with these genes are also ligated into the expression vector 3' of the sequence desired to be expressed to provide polyadenylation of the mRNA and termination.

[0071] Other suitable promoters that have the additional advantage of transcription controlled by growth conditions include the promoter region for alcohol dehydrogenase 2, isocytichrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization.

[0072] In addition to microorganisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is operable, whether from vertebrate or invertebrate culture. In addition to mammalian cells, these include insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus); and plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing one or more coding sequences.

[0073] In a useful insect system, *Autographica californica nuclear polyhedrosis* virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The isolated nucleic acid coding sequences are cloned into non-essential regions (for example the polyhedron gene) of the virus and placed under control of an AcNPV promoter (for example, the polyhedron promoter). Successful insertion of the coding sequences results in the inactivation of the polyhedron gene and production of non-occluded recombinant virus (i.e., virus lacking the proteinaceous coat coded for by the polyhedron gene). These recombinant viruses are then used to infect *Spodoptera frugiperda* cells in which the inserted gene is expressed (e.g., U.S. Patent No. 4,215,051).

[0074] Examples of useful mammalian host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, W138, BHK, COS-7, 293, HepG2, NIH3T3, RIN and MDCK cell lines. In addition, a host cell may be chosen that modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the encoded protein.

[0075] Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. Expression vectors for use in mammalian cells optionally include an origin of replication (as necessary), a promoter located in front of the gene to be expressed, along with any necessary ribosome binding sites, RNA splice sites, polyadenylation site, and transcriptional terminator sequences. The origin of replication may be provided either by construction of the vector to include an exogenous origin, such as may be derived from SV40 or other viral (e.g., Polyoma, Adeno, VSV, BPV) source, or may be provided by the host cell chromosomal replication mechanism. If the vector is integrated into the host cell chromosome, the latter is often sufficient.

[0076] The promoters may be derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter). Further, it is also possible, and may be desirable, to utilize promoter or control sequences normally associated with the desired gene sequence, provided such control sequences are compatible with the host cell systems.

[0077] A number of viral based expression systems may be utilized, for example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40

(SV40). The early and late promoters of SV40 virus are useful because both are obtained easily from the virus as a fragment that also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250 bp sequence extending from the *HmdIII* site toward the *BglI* site located in the viral origin of replication.

[0078] In cases where an adenovirus is used as an expression vector, the coding sequences may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by *in vitro* or *in vivo* recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing proteins in infected hosts.

[0079] Specific initiation signals may also be required for efficient translation of the claimed isolated nucleic acid coding sequences. These signals include the ATG initiation codon and adjacent sequences. Exogenous translational control signals, including the ATG initiation codon, may additionally need to be provided. One of ordinary skill in the art would readily be capable of determining this need and providing the necessary signals. It is well known that the initiation codon must be in-frame (or in-phase) with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements or transcription terminators.

[0080] In eukaryotic expression, one will also typically desire to incorporate into the transcriptional unit an appropriate polyadenylation site if one was not contained within the original cloned segment. Typically, the poly A addition site is placed about 30 to 2000 nucleotides downstream of the termination site of the protein at a position prior to transcription termination.

[0081] For long-term, high-yield production of recombinant proteins, stable expression is optionally achieved. For example, cell lines that stably express constructs encoding proteins may be engineered. Rather than using expression vectors that contain viral origins of replication, host cells can be transformed with vectors controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched medium, and then are switched to a selective medium. The

selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci, which in turn can be cloned and expanded into cell lines.

[0082] A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase, hypoxanthine-guanine phosphoribosyltransferase and adenine phosphoribosyltransferase genes, in tk^- , $hgprt^-$ or $aprt^-$ cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for: dhfr, which confers resistance to methotrexate; gpt, which confers resistance to mycophenolic acid; neo, which confers resistance to the aminoglycoside G-418; and hygromycin. It is appreciated that numerous other selection systems are known in the art that are similarly operable in the present invention.

[0083] It is contemplated that the isolated nucleic acids of the disclosure may be "overexpressed", i.e., expressed in increased levels relative to its natural expression in cells of its indigenous host, or even relative to the expression of other nucleic acids or proteins in a recombinant host cell. Overexpression may be assessed by a variety of methods, including radio-labeling and/or protein purification. However, simple and direct methods are may be used, for example, those involving SDS/PAGE and protein staining or immunoblotting, followed by quantitative analyses, such as densitometric scanning of the resultant gel or blot. A specific increase in the level of the recombinant nucleic acid or protein in comparison to the level in natural host cells is indicative of overexpression, as is a relative abundance of the specific protein in relation to the other proteins produced by the host cell and, e.g., visible on a gel.

[0084] Other methods, primers, isolation techniques, sequencing techniques, and characterization techniques are known to those of skill in the art and are similarly operable herein. Illustratively, one can reconstitute SHIBV viruses *de novo* from isolated genes such as by assembly of virus particles with captured genes illustratively by the techniques of or modifications of Gonzalez, SA, and Affranchino, JL, J. Gen. Virol., 1995; 76:2357-2360, the contents of which are incorporated herein by reference.

[0085] Further aspects of the present disclosure concern the purification, and in particular embodiments, the substantial purification, of an encoded protein or peptide. The term "purified" or "isolated" protein or peptide as used herein, is intended to refer to a composition, isolatable from other components, wherein the nucleic acid, protein or peptide is purified to any degree relative to its naturally-obtainable state, i.e., relative to its purity within a cell, relative to its purity within a virion, or relative to its purity within an infective organism. An isolated nucleic acid,

protein, or peptide also refers to a nucleic acid, protein or peptide, free from the environment in which it may naturally occur.

[0086] Generally, "purified" or "isolated" will refer to a nucleic acid, protein or peptide composition that has been subjected to fractionation to remove various other components, and which composition substantially retains its expressed biological activity, if any. Where the term "substantially" purified is used, this designation will refer to a composition in which the nucleic acid, protein or peptide forms the major component of the composition, such as constituting about 50% or more of the nucleic acids or proteins in the composition.

[0087] Various methods for quantifying the degree of purification of the nucleic acid, protein or peptide will be known to those of skill in the art. These include, for example, determining the specific activity of an active fraction, or assessing the number of polypeptides within a fraction by SDS/PAGE analysis. An optional method for assessing the purity of a fraction is to calculate the specific activity of the fraction, to compare it to the specific activity of the initial extract, and to thus calculate the degree of purity, herein assessed by a "-fold purification number." The actual units used to represent the amount of activity will, of course, be dependent upon the particular assay technique chosen to follow the purification and whether or not the expressed nucleic acid, protein or peptide exhibits a detectable activity.

[0088] Various techniques suitable for use in protein purification will be well known to those of skill in the art. These include, for example: precipitation with ammonium sulphate, polyethylene glycol, antibodies and the like or by heat denaturation, followed by centrifugation; chromatography steps such as ion exchange, gel filtration, reverse phase, hydroxylapatite and affinity chromatography; isoelectric focusing; gel electrophoresis; and combinations of such and other techniques. As is generally known in the art, it is believed that the order of conducting the various purification steps may be changed, or that certain steps may be omitted, and still result in a suitable method for the preparation of a substantially purified protein or peptide.

[0089] There is no general requirement that the nucleic acid, protein or peptide always be provided in its most purified state. Indeed, it is contemplated that less substantially purified products will have utility in certain embodiments. Partial purification may be accomplished by using fewer purification steps in combination, or by utilizing different forms of the same general purification scheme. For example, it is appreciated that a cation-exchange column chromatography performed utilizing an HPLC apparatus will generally result in a greater -fold purification than the same technique utilizing a low pressure chromatography system. Methods exhibiting a lower degree of relative purification may have advantages in total recovery of

product, or in maintaining the activity of an expressed protein. A protein or nucleic acid is optionally substantially purified to a purity of 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more, or any value or range therebetween.

[0090] It is known that the migration of a polypeptide can vary, sometimes significantly, with different conditions of SDS/PAGE (Capaldi et al., *Biochem. Biophys. Res. Comm.*, 76:425, 1977). It will therefore be appreciated that under differing electrophoresis conditions, the apparent molecular weights of purified or partially purified expression products may vary.

[0091] The inventive method also illustratively includes isolation of glycoprotein from a host cell or host cell medium. Methods of protein isolation illustratively include column chromatography, affinity chromatography, gel electrophoresis, filtration, or other methods known in the art. In some embodiments, glycoprotein is expressed with a tag operable for affinity purification. A tag is optionally a 6x His tag. A 6x His tagged inventive protein is illustratively purified by Ni-NTA column chromatography or using an anti-6x His tag antibody fused to a solid support. (Geneway Biotech, San Diego, CA) Other tags and purification systems are similarly operable.

[0092] It is appreciated that an inventive protein is optionally not tagged. In this embodiment and other embodiments purification is optionally achieved by methods known in the art illustratively including ion-exchange chromatography, affinity chromatography using anti-glycoprotein antibodies, precipitation with salt such as ammonium sulfate, streptomycin sulfate, or protamine sulfate, reverse phase chromatography, size exclusion chromatography such as gel exclusion chromatography, HPLC, immobilized metal chelate chromatography, or other methods known in the art. One of skill in the art may select the most appropriate isolation and purification techniques without departing from the scope of this invention.

[0093] Inventive nucleic acids, proteins or fragments thereof are optionally chemically synthesized. Methods of protein chemical synthesis have produced proteins greater than 600 amino acids in length with or without the inclusion of modifications such as glycosylation and phosphorylation. Methods of chemical protein and peptide synthesis illustratively include solid phase protein chemical synthesis. Illustrative methods of chemical protein synthesis are reviewed by Miranda, LP, *Peptide Science*, 2000, 55:217-26 and Kochendoerfer GG, *Curr Opin Drug Discov Devel.* 2001; 4(2):205-14, the contents of which are incorporated herein by reference.

[0094] Proteins of this invention are optionally be characterized by immunological measurements including, without limitation, western blot, macromolecular mass determinations

by biophysical determinations, SDS-PAGE/staining, HPLC and the like, antibody recognition assays, cell viability assays, apoptosis assays, and assays to infer immune protection or immune pathology by adoptive transfer of cells, proteins or antibodies.

[0095] Vaccines and methods for their use to induce active immunity and protection against lyssavirus induced illness in a subject are provided according to the present invention.

[0096] The term "vaccine composition" is used herein to refers to a composition including a biological agent (e.g. protein or nucleic acid) capable of inducing an immune response in a subject inoculated with the vaccine composition. In particular embodiments, the biological agent is a live attenuated and/or inactive SHIBV. In further embodiments, the biological agent is an antigenic portion of a SHIBV.

[0097] A SHIBV included in a vaccine composition according to the present invention is a live attenuated SHIBV or an inactivated SHIBV. The choice of live attenuated SHIBV or inactivated SHIBV depends on factors such as route of vaccine composition administration. The isolated SHIBV is optionally treated to inactivate or attenuate the lyssavirus. Thus, in particular embodiments a vaccine for *lyssavirus* includes a live attenuated SHBIV or an inactivated human SHBIV.

[0098] The term "live attenuated SHIBV" refers to a SHIBV having the ability to infect an appropriate host or host cell and replicate and the term is used to distinguish an "inactivated" SHIBV. The term "live attenuated SHIBV" refers to a SHIBV characterized by substantially diminished virulence compared to wild type SHIBV. The term "virulence" is used to describe the degree of pathogenicity of a SHIBV to a host cell or a host organism. Virulence is determined using any of various assays recognized in the art. For example, virulence may be assessed by exposing cultured host cells to an attenuated SHIBV and determining the number of cells which display a pathogenic response and/or the severity of pathogenic response elicited. Diminished virulence is present where an attenuated SHIBV has decreased capability to cause one or more pathogenic effects in a host cell and/or host organism.

[0099] The term "inactivated" SHIBV is used herein to refer to SHIBV that has been killed and which is therefore capable of neither replication nor infection of a host cell or host organism.

[00100] Inactivation is achieved by any of various techniques illustratively including inactivation using one or more chemical agents, thermal inactivation and/or UV light inactivation.

[00101] Chemical agents used to inactivate a SHIBV are known in the art and include such agents as ethyleneimines such as binary ethyleneimine; cross-linking aldehydes such as

formaldehyde and glutaraldehyde; proteases illustratively including pronase, trypsin and/or chymotrypsin; and detergents such as octylphenol ethoxylates and alkyl trimethylammonium salts. SHIBV may be inactivated by treatment with a base, for example by incubation of the lyssavirus at a pH above pH 10.0.

5 **[00102]** Thermal inactivation may be achieved by heating at temperatures above 50° centigrade, for example.

[00103] Inactivation is assessed by techniques standard in the art, illustratively including sampling virus at various times during an inactivation procedure and observing cytopathic effects or infectivity of a sample on suitable cells, such as mouse neuroblastoma (MNA) cells.

10 **[00104]** It is appreciated that, in addition to live attenuated and inactivated SHIBV, an antigenic portion of a SHIBV is optionally included in a vaccine composition of the present invention. Thus, for example, a SHIBV-derived protein or peptide capable of inducing an immunological response in a subject is considered within the scope of the present invention.

[00105] Vaccine compositions are provided according to embodiments of the present
15 invention which include one or more SHIBV polypeptides and/or an immunogenic fragment of one or more SHIBV polypeptides. In particular embodiments of an inventive vaccine composition, a SHIBV polypeptide, a derivative thereof, and/or an immunogenic fragment thereof is included.

[00106] Accordingly, the present invention provides a virus including an nucleoprotein
20 having SEQ ID NO: 2 or having an amino acid sequence that is greater than 80%, is greater than 81%, greater than 82%, greater than 83%, greater than 84%, greater than 85%, greater than 86%, greater than 87%, greater than 88%, greater than 89%, greater than 90%, greater than 91%, greater than 92%, greater than 93%, greater than 94%, greater than 95%, greater than 96%, greater than 97%, greater than 98% or greater than 99% identical to SEQ ID NO: 2.

25 **[00107]** The present invention provides a virus including an phosphoprotein having SEQ ID NO: 3 or having an amino acid sequence that is greater than 80%, is greater than 81%, greater than 82%, greater than 83%, greater than 84%, greater than 85%, greater than 86%, greater than 87%, greater than 88%, greater than 89%, greater than 90%, greater than 91%, greater than 92%, greater than 93%, greater than 94%, greater than 95%, greater than 96%, greater than 97%,
30 greater than 98% or greater than 99% identical to SEQ ID NO: 3.

[00108] The present invention provides a virus including an matrix protein having SEQ ID NO: 4 or having an amino acid sequence that is greater than 80%, is greater than 81%, greater than 82%, greater than 83%, greater than 84%, greater than 85%, greater than 86%, greater than

87%, greater than 88%, greater than 89%, greater than 90%, greater than 91%, greater than 92%, greater than 93%, greater than 94%, greater than 95%, greater than 96%, greater than 97%, greater than 98% or greater than 99% identical to SEQ ID NO: 4.

[00109] The present invention provides a virus including an glycoprotein having SEQ ID NO: 5 or having an amino acid sequence that is greater than 80%, is greater than 81%, greater than 82%, greater than 83%, greater than 84%, greater than 85%, greater than 86%, greater than 87%, greater than 88%, greater than 89%, greater than 90%, greater than 91%, greater than 92%, greater than 93%, greater than 94%, greater than 95%, greater than 96%, greater than 97%, greater than 98% or greater than 99% identical to SEQ ID NO: 5.

[00110] The present invention provides a virus including an polymerase protein having SEQ ID NO: 6 or having an amino acid sequence that is greater than 80%, is greater than 81%, greater than 82%, greater than 83%, greater than 84%, greater than 85%, greater than 86%, greater than 87%, greater than 88%, greater than 89%, greater than 90%, greater than 91%, greater than 92%, greater than 93%, greater than 94%, greater than 95%, greater than 96%, greater than 97%, greater than 98% or greater than 99% identical to SEQ ID NO: 6.

[00111] In particular embodiments of the invention, a SHIBV for inclusion in a vaccine composition of the present invention is prepared by standard methods typically used for preparation of live or inactivated SHIBV. For example, generally a compatible cell type is inoculated with a SHIBV and the cells are maintained under conditions which allow for viral replication and production of infectious particles.

[00112] Inventive viral proteins of the present invention are optionally modified to increase their immunogenicity. In a non-limiting example, the antigen may be coupled to chemical compounds or immunogenic carriers, provided that the coupling does not interfere with the desired biological activity of either the antigen or the carrier. For a review of some general considerations in coupling strategies, see Antibodies , A Laboratory Manual, Cold Spring Harbor Laboratory, ed. E. Harlow and D. Lane (1988). Useful immunogenic carriers known in the art include, without limitation: keyhole limpet hemocyanin (KLH); bovine serum albumin (BSA); ovalbumin; PPD (purified protein derivative of tuberculin); red blood cells; tetanus toxoid; cholera toxoid; agarose beads; activated carbon; or bentonite. Useful chemical compounds for coupling include, without limitation, dinitrophenol groups and arsonilic acid.

[00113] SHIBV particles are harvested, typically from cell culture supernatant for inclusion in a vaccine composition. The SHIBV particles may be isolated from the cell culture

supernatant, for example by filtration and/or centrifugation. The isolated SHIBV particles are optionally lyophilized, such as for later resuspension in a pharmaceutically acceptable carrier.

[00114] The term "pharmaceutically acceptable carrier" refers to a carrier which is substantially non-toxic to a subject and substantially inert to the lyssavirus included in a vaccine composition. A pharmaceutically acceptable carrier is a solid, liquid or gel in form and is typically sterile and pyrogen free.

[00115] A vaccine composition of the present invention may be in any form suitable for administration to a subject.

[00116] A vaccine composition is administered by any suitable route of administration including oral and parenteral such as intravenous, intradermal, intramuscular, mucosal, nasal, or subcutaneous routes of administration.

[00117] For example, a vaccine composition for parenteral administration may be formulated as an injectable liquid including a SHIBV, SHBIV protein or nucleic acid, or fragment thereof, and a pharmaceutically acceptable carrier. Examples of suitable aqueous and nonaqueous carriers include water, ethanol, polyols such as propylene glycol, polyethylene glycol, glycerol, and the like, suitable mixtures thereof; vegetable oils such as olive oil; and injectable organic esters such as ethyloleate. Proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of a desirable particle size in the case of dispersions, and/or by the use of a surfactant, such as sodium lauryl sulfate. A stabilizer is optionally included such as, for example, sucrose, EDTA, EGTA, and an antioxidant.

[00118] A solid dosage form for administration or for suspension in a liquid prior to administration illustratively includes capsules, tablets, powders, and granules. In such solid dosage forms, a SHIBV, SHBIV protein or nucleic acid, or fragment thereof, is admixed with at least one carrier illustratively including a buffer such as, for example, sodium citrate or an alkali metal phosphate illustratively including sodium phosphates, potassium phosphates and calcium phosphates; a filler such as, for example, starch, lactose, sucrose, glucose, mannitol, and silicic acid; a binder such as, for example, carboxymethylcellulose, alginates, gelatin, polyvinylpyrrolidone, sucrose, and acacia; a humectant such as, for example, glycerol; a disintegrating agent such as, for example, agar-agar, calcium carbonate, plant starches such as potato or tapioca starch, alginic acid, certain complex silicates, and sodium carbonate; a solution retarder such as, for example, paraffin; an absorption accelerator such as, for example, a quaternary ammonium compound; a wetting agent such as, for example, cetyl alcohol, glycerol monostearate, and a glycol; an adsorbent such as, for example, kaolin and bentonite; a lubricant

such as, for example, talc, calcium stearate, magnesium stearate, a solid polyethylene glycol or sodium lauryl sulfate; a preservative such as an antibacterial agent and an antifungal agent, including for example, sorbic acid, gentamycin and phenol; and a stabilizer such as, for example, sucrose, EDTA, EGTA, and an antioxidant.

5 **[00119]** Solid dosage forms optionally include a coating such as an enteric coating. The enteric coating is typically a polymeric material. Enteric coating materials optionally have the characteristics of being bioerodible, gradually hydrolyzable and/or gradually water-soluble polymers. The amount of coating material applied to a solid dosage generally dictates the time interval between ingestion and drug release. A coating is applied having a thickness such that
10 the entire coating does not dissolve in the gastrointestinal fluids at pH below 3 associated with stomach acids, yet dissolves above pH 3 in the small intestine environment. It is expected that any anionic polymer exhibiting a pH-dependent solubility profile is readily used as an enteric coating in the practice of the present invention to achieve delivery of the active agent to the lower gastrointestinal tract. The selection of the specific enteric coating material depends on
15 properties such as resistance to disintegration in the stomach; impermeability to gastric fluids and active agent diffusion while in the stomach; ability to dissipate at the target intestine site; physical and chemical stability during storage; non-toxicity; and ease of application.

[00120] Suitable enteric coating materials illustratively include cellulosic polymers such as hydroxypropyl cellulose, hydroxyethyl cellulose, hydroxypropyl methyl cellulose, methyl
20 cellulose, ethyl cellulose, cellulose acetate, cellulose acetate phthalate, cellulose acetate trimellitate, hydroxypropylmethyl cellulose phthalate, hydroxypropylmethyl cellulose succinate and carboxymethylcellulose sodium; acrylic acid polymers and copolymers, preferably formed from acrylic acid, methacrylic acid, methyl acrylate, ammonium methylacrylate, ethyl acrylate, methyl methacrylate and/or ethyl; vinyl polymers and copolymers such as polyvinyl pyrrolidone,
25 polyvinyl acetate, polyvinylacetate phthalate, vinylacetate crotonic acid copolymer, and ethylene-vinyl acetate copolymers; shellac; and combinations thereof. A particular enteric coating material includes acrylic acid polymers and copolymers described for example U.S. Patent No. 6,136,345.

[00121] The enteric coating optionally contains a plasticizer to prevent the formation of pores and cracks that allow the penetration of the gastric fluids into the solid dosage form. Suitable
30 plasticizers illustratively include, triethyl citrate (Citroflex 2), triacetin (glyceryl triacetate), acetyl triethyl citrate (Citroflex A2), Carbowax 400 (polyethylene glycol 400), diethyl phthalate, tributyl citrate, acetylated monoglycerides, glycerol, fatty acid esters, propylene glycol, and

dibutyl phthalate. In particular, a coating composed of an anionic carboxylic acrylic polymer typically contains approximately 10% to 25% by weight of a plasticizer, particularly dibutyl phthalate, polyethylene glycol, triethyl citrate and triacetin. The coating can also contain other coating excipients such as detackifiers, antifoaming agents, lubricants (e.g., magnesium stearate),
5 and stabilizers (e.g. hydroxypropylcellulose, acids or bases) to solubilize or disperse the coating material, and to improve coating performance and the coated product.

[00122] Liquid dosage forms for oral administration include lyssavirus and a pharmaceutically acceptable carrier formulated as an emulsion, solution, suspension, syrup, or elixir. A liquid dosage form of a vaccine composition of the present invention may include a
10 wetting agent, an emulsifying agent, a suspending agent, a sweetener, a flavoring, or a perfuming agent.

[00123] Detailed information concerning customary ingredients, equipment and processes for preparing dosage forms is found in *Pharmaceutical Dosage Forms: Tablets*, eds. H. A. Lieberman et al., New York: Marcel Dekker, Inc., 1989; and in L.V. Allen, Jr. et al., *Ansel's
15 Pharmaceutical Dosage Forms and Drug Delivery Systems*, 8th Ed., Philadelphia, PA: Lippincott, Williams & Wilkins, 2004, throughout and in chapter 16; A. R. Gennaro, *Remington: The Science and Practice of Pharmacy*, Lippincott Williams & Wilkins, 21st ed., 2005, particularly chapter 89; and J. G. Hardman et al., *Goodman & Gilman's The Pharmacological Basis of Therapeutics*, McGraw-Hill Professional, 10th ed., 2001.

[00124] An adjuvant is optionally included in a virus composition according to embodiments of the present invention. Adjuvants are known in the art and illustratively include Freund's adjuvant, aluminum hydroxide, aluminum phosphate, aluminum oxide, saponin, dextrans such as DEAE-dextran, vegetable oils such as peanut oil, olive oil, and/or vitamin E acetate, mineral oil, bacterial lipopolysaccharides, peptidoglycans, and proteoglycans.

[00125] The term "subject" is used herein to refer to a human, non-human animals, illustratively including other primates, cows, horses, sheep, goats, pigs, dogs, cats, birds, poultry, and rodents such as mice or rats.

[00126] Viral proteins of the present invention may also be used in the form of pharmaceutically acceptable salts. Suitable acids and bases which are capable of forming salts
30 with the proteins of the present invention are well known to those of skill in the art, and include inorganic and organic acids and bases.

[00127] The proteins and nucleic acid sequences or anti-sense sequences of the invention, alone or in combination with other antigens, antibodies, nucleic acid sequences or anti-sense

sequences may further be used in diagnostic methods, therapeutic compositions and in methods for treating humans and/or animals with disease or at risk of disease or infection. For example, one such therapeutic composition may be formulated to contain a carrier or diluent and one or more Shimoni bat virus proteins or protein fragments of the invention. For diagnostic purposes, as an example the original or synthesized G gene is optionally inserted into foreign vectors (for example, a pseudotype lentiviral vector), and further used for specific virus-neutralization diagnostic assays. Suitable pharmaceutically acceptable carriers facilitate administration of the proteins but are physiologically inert and/or nonharmful.

[00128] Optionally, the inventive composition may also contain conventional pharmaceutical ingredients, such as preservatives, or chemical stabilizers. Suitable ingredients operable herein include, for example, casamino acids, sucrose, gelatin, phenol red, N-Z amine, monopotassium diphosphate, lactose, lactalbumin hydrolysate, and dried milk.

[00129] In particular embodiments, a vaccine composition including a SHIBV, and/or SHIBV polypeptides and/or an immunogenic fragment of one or more SHIBV polypeptides, stimulates generation of neutralizing antibodies to a SHIBV or other immunological response in a subject. Methods of inducing an immunological response against a SHIBV-mediated disease in a subject are provided according to embodiments of the present invention which include administering a therapeutic amount of a vaccine composition including at least one SHIBV, SHIBV protein, or fragment thereof.

[00130] The phrase "therapeutically effective amount" is used herein to refer to an amount effective to induce an immunological response sufficient to prevent or ameliorate signs or symptoms of a *lyssavirus-mediated* disease. Induction of an immunological response in a subject can be determined by any of various techniques known in the art, illustratively including detection of *anti-lyssavirus* antibodies, measurement of *anti-lyssavirus* antibody titer and/or lymphocyte proliferation assay. Illustrative methods for detection of *anti-lyssavirus* antibodies are illustrated by Hanlon, CA., et al., *Virus Res.*, 2005; 111(1):44-54, the contents of which are incorporated herein by reference. Signs and symptoms of *lyssavirus-mediated* disease may be monitored to detect induction of an immunological response to administration of a vaccine composition of the present invention in a subject. An immunological response is illustratively a reduction of clinical signs and symptoms of *lyssavirus-mediated* disease. An immunological response is illustratively, development of anti- *lyssavirus* antibodies, activation of T-cells, B-cells, or other immune cells following administration of an inventive composition, or other immune responses known in the art.

[00131] In some embodiments, a method of inducing an immunological response against a *lyssavirus-mediated* disease in a subject includes administering 10^4 to 10^8 ffu of live attenuated vaccine or 1 to 25 micrograms of inactivated virus in a typical vaccine composition.

[00132] In some embodiments, a method of inducing an immunological response against a *lyssavirus-mediated* disease in a subject includes administering a therapeutically effective amount of a vaccine composition including a SHIBV, polypeptide fragments thereof, derivatives thereof, or combinations thereof.

[00133] Administration of a vaccine composition according to a method of the present invention includes administration of one or more doses of a vaccine composition to a subject at one time in particular embodiments. Alternatively, two or more doses of a vaccine composition are administered at time intervals of days, weeks, or years. A suitable schedule for administration of vaccine composition doses depends on several factors including age and health status of the subject, type of vaccine composition used and route of administration, for example. One of skill in the art is able to readily determine a dose and schedule of administration to be administered to a particular subject.

[00134] Also provided is a method for detecting the presence of a lyssavirus virus or a nucleic acid molecule derived from a lyssavirus in a biological sample including contacting a sample with an agent that selectively binds to the a nucleic acid molecule with a sequence that represents at least a portion of SEQ ID NO: 1; or a protein sequence encoded by at least a portion of SEQ ID NO: 1, and detecting whether the agent binds to the virus or the nucleic acid molecule derived therefrom in the sample. In some embodiments the agent has a nucleic acid sequence that hybridizes under stringent conditions with at least a portion of SEQ ID NO: 1, but optionally, has 100% sequence identity or less than 100% sequence identity. In some embodiments, the agent is a nucleic acid molecule with a nucleotide sequence having between 4 and 6600 contiguous nucleotides of the nucleotide sequence of SEQ ID NO: 1, or a complement thereof.

[00135] An agent optionally selectively binds a protein with a sequence of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, or combinations thereof.

[00136] The term "selectively binds" describes binding to a target nucleic acid sequence or amino acid sequence with less than 1% cross reactivity toward another nucleic acid sequence or amino acid sequence present in a sample.

[00137] As used herein, the term "hybridizes under stringent conditions" describes conditions for hybridization and washing under which nucleotide sequences having at least 30%, 35%,

40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, or 95% identity to each other typically remain hybridized to each other. Such hybridization conditions are described in, for example but not limited to, Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1 6.3.6.; Basic Methods in Molecular Biology, Elsevier Science Publishing Co., Inc., N.Y. (1986), pp.75 78, and 84 87; and Molecular Cloning, Cold Spring Harbor Laboratory, N.Y. (1982), pp.387 389, and are well known to those skilled in the art. A non-limiting example of stringent hybridization conditions is hybridization in 6x sodium chloride/sodium citrate (SSC), 0.5% SDS at about 68°C followed by one or more washes in 2xSSC, 0.5% SDS at room temperature. Another non-limiting example of stringent hybridization conditions is hybridization in 6x SSC at about 45°C followed by one or more washes in 0.2x SSC, 0.1% SDS at 50 to 65 °C.

[00138] An agent is illustratively any molecule that will selectively bind at least a portion of a nucleic acid sequence with the sequence of SEQ ID NO: 1 or a protein encoded by at least a portion of SEQ ID NO: 1, or derivatives thereof. Illustrative examples of an agent include an antibody, other proteinaceous molecules, an aptamer, a nucleotide sequence, other molecules such as low molecular weight compositions with a molecular weight below 2000 Daltons, or other nucleic acid or protein binding compositions known in the art.

[00139] An agent is illustratively an antibody. Such antibodies can be polyclonal or monoclonal. An intact antibody, a fragment thereof (e.g., Fab or F(ab')₂), or an engineered variant thereof (e.g., sFv) can also be used. Such antibodies can be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. Methods of producing antibodies directed to specific sequences of protein or nucleic acid are known in the art. Optional methods of producing and screening for antibodies that will act as an agent are described by Birch, JR, and Racher, AJ, *Advanced Drug Delivery Reviews*, 2006; 58:671- 685.

[00140] A biological sample is any sample derived from a biological source including an animal, plant, tissue, or cell. A biological sample is illustratively tissue derived from brain such as the brain stem or cerebellum, other neuronal tissue, kidney, heart, or other tissue. In some embodiments, a biological sample is blood, plasma, serum, urine, feces, saliva, nasal secretions, lung aspirate, cerebrospinal fluid, or skin.

[00141] A process of the invention illustratively includes detecting whether the agent binds to the virus or the nucleic acid molecule derived therefrom in the sample. Detection is by any method known in the art suitable for detecting the binding of an agent to a nucleic acid molecule or a protein. Illustrative examples include PCR related techniques such as PCR, RT-PCR, real-time PCR (qRT-PCR), hybridization of labeled agents where an agent is illustratively labeled

with a radioactive or fluorescent molecule, mass spectrometry, electrophoresis or other technique known in the art. Illustrative examples of detecting the binding of an agent to a protein sequence include gel electrophoresis (e.g. SDS-PAGE and others), enzyme linked immunosorbent assay (ELISA), mass spectrometry, gradient sedimentation, fluorescence, NMR, liquid chromatography, and other methods known in the art. Illustrative examples of methods suitable for detection are described in references listed herein.

[00142] Various aspects of the present invention are illustrated by the following non-limiting examples. The examples are for illustrative purposes and are not a limitation on any practice of the present invention. It will be understood that variations and modifications can be made without departing from the spirit and scope of the invention.

Example 1:

[00143] Sequencing the SHIBV genome. Total RNA was extracted from bat brain homogenates using TRIzol (Invitrogen Corp. Carlsbad, CA) and subjected to RT-PCR with subsequent direct sequencing on the AB13730 automated sequencer (Applied Biosystems, Foster City, CA). Genome termini are circularized by RNA ligation, amplified by RT-PCR, cloned and sequenced as described previously (Kuzmin et al., *J. Clin. Microbiol.*, 2008; 46(4):1451-1461). Each DNA strand of a given PCR product is sequenced at least twice. The sequence assembly, alignment and consensus sequence generation, as well as DNA translation into deduced amino acids are performed in BioEdit software (Hall, T.A., *Nucl. Acids. Symp. Ser.*, 1999; 41:95-98).

[00144] The length of the complete SHIBV genome is 12045 nucleotides (GenBank Accession No. GU170201). The genome consists of 5 structural genes, common for all lyssaviruses, and non-coding regions that are most similar to those described in the phylogroup II lyssaviruses. Characteristics and alignments of particular genes of SHIBV are illustrated in Kuzmin, I, et al., *Virus Res.*, 2010;149(2):197-210.

Example 2:

[00145] Alignment with known lyssavirus sequences. A set of representative lyssaviruses is shown in Table 1. For the phylogroup II lyssaviruses, all available sequences are included, whereas for the phylogroup I lyssaviruses, a selection is made to cover the intrinsic diversity of each species. Multiple alignments for each viral gene, and for the alignment of concatenated coding regions of the N, P, M, G and L genes (both nucleotide and deduced amino acid sequences), are produced using the ClustalX program (Jeanmougin et al., 1998). The identity

values are calculated in BioEdit. Neighbor joining (NJ) phylogenetic analysis is performed in MEGA program (Kumar et al., 2001), using p-distances, Kimura-2 parameters and Maximum composite likelihood models, for 1000 bootstrap replicates. Bayesian analysis (BI) is performed using BEAST software (Drummond and Rambaut, 2007), with the general time-reversible model incorporating both invariant sites and a gamma distribution (GTR+I+G). Two simultaneous analyses, each with four Markov chains, are run for 1,000,000 generations and sampled every 1,000 generations. Trees generated prior to the stabilization of likelihood scores are discarded (burning = 250). The remaining trees are used to build a 50% majority rule consensus tree. Maximum likelihood (ML) analysis is performed using the PHYLIP package (Felsenstein, 1993) for 100 bootstrap replicates. Nucleotide substitution models used transition/transversion ratios varying from 2 to 4, with empirical base frequencies, and a gamma distribution of rate variations among sites. The gamma parameter for each alignment was determined using PAUP* (Swofford, 2003).

[00146] Table 1:

Virus species and code	Species isolated from	Country	GenBank No	accession
RABV:				
SAD B19	Laboratory strain			M31046
PV	Laboratory strain			M13215
9147FRA	Red fox	France		EU293115
AY956319	Human (ex a dog)	Germany (ex a India)	(ex	AY956319
8764THA	Human (ex a dog)	Thailand		EU293111
9704ARG	Bat <i>Tadarida brasiliensis</i>	Argentina		EU293116
SHBRV-18	Human (ex bat <i>Lasionycteris noctivagans</i>)	USA		AY705373

LBV:

LBVAFR1999		France (ex Togo	EF547447;
	Bat	<i>Rousettus</i> or Egypt)	EF547418;
		<i>aegyptiacus</i>	EF547445;
			EF547432
LBVSEN1985		Senegal	EF547448;
	Bat	<i>Eidolon</i>	EF547419;
		<i>helvum</i>	EF547446;
			EF547433;
(0406SEN)			EU293108
KE131	Bat	<i>Eidolon</i> Kenya	EU259198
		<i>helvum</i>	
KE576	Bat	<i>Rousettus</i> Kenya	GU170202
		<i>aegyptiacus</i>	
LBVNIG1956		Nigeria	EF547459;
	Bat	<i>Eidolon</i>	EF547407;
		<i>helvum</i>	EF547444;
			EF547431;
(8619NGA)			EU293110
LBVCAR1974	Bat	Central African	EF547449;
		Republic	EF547417;
		<i>Micropteropus</i>	EF547443;
		<i>pussilus</i>	EF547430
LBVZIM1986		Zimbabwe	EF547450;
	Cat		EF547416;
			EF547442;
			EF547429
LagSA2004	Bat	South Africa	EF547458;
		<i>Epomophorus</i>	EF547415;
		<i>wahlbergi</i>	EF547440;
			EF547428
LagSA2003	Bat	South Africa	EF547451;
		<i>Epomophorus</i>	EF547413;
		<i>wahlbergi</i>	EF547434;

			EF547421
Mongoose2004		South Africa	EF547453;
	Mongoose		EF547409;
			EF547438;
			EF547423
LBVSA1980		South Africa	EF547454;
	Bat		EF547411;
	<i>Epomophorus</i>		EF547441;
	<i>wahlbergi</i>		EF547424
LBVSA1982		South Africa	EF547455;
	Bat		EF547410;
	<i>Epomophorus</i>		EF547439;
	<i>wahlbergi</i>		EF547425
MOKV:			
Y09762	Cat	Zimbabwe	Y09762
86100CAM	Shrew	Cameroon	EU293117
86101RCA	A rodent	Central African Republic	EU293118
DUVV:			
94286SA	Bat <i>Miniopterus</i>	South Africa	EU293120
	sp. (?)		
86132SA	Human (ex a bat)	South Africa	EU293119
EBLV-1:			
8918FRA	Bat <i>Eptesicus</i>	France	EU293112
	<i>serotinus</i>		
03002FRA	Bat <i>Eptesicus</i>	France	EU293109
	<i>serotinus</i>		
RV9	Bat <i>Eptesicus</i>	Germany	EF157976
	<i>serotinus</i>		
EBLV2:			
RV1333	Human (ex a bat)	United Kingdom	EF157977

9018HOL	Bat	<i>Myotis</i>	The Netherlands	EU293114
		<i>dasycneme</i>		
AF418014	Human (ex a bat)		Australia	AF418014
AF081020	Bat	<i>Saccolaimus</i>	Australia	AF081020
		<i>albiventris</i>		
NC_003243	Bat	<i>Saccolaimus</i>	Australia	NC_003243
		<i>albiventris</i>		
ΛRAV	Bat	<i>Myotis blythi</i>	Kyrgyzstan	EF614259
KHUV	Bat	<i>Myotis</i>	Tajikistan	EF614261
		<i>mystacinus</i>		
IRKV	Bat	<i>Murina</i>	Russia	EF614260
		<i>leucogaster</i>		
WCBV	Bat	<i>Miniopterus</i>	Russia	EF614258
		<i>schreibersi</i>		

[00147] The aligned sequences of N+P+M+G+L genes of available lyssaviruses (total length of the alignment 10922 nucleotides), as well as the deduced protein sequences (total length of the alignment 3639 amino acids) are assembled and used for comparison to a concatenated SHIBV from the deduced amino acid sequences obtained from the genome sequence of Example 1. Characteristics and alignments of particular genes of SHIBV as well as the concatenated sequences are illustrated in Kuzmin, I, et al., *Virus Res.*, 2010;149(2):197-210.

Example 3:

[00148] Monoclonal antibody typing of SHIBV. The newly isolated virus and representatives of all described lineages of LBV and MOKV are inoculated intracranially into suckling or 3-week-old mice. When clinical signs of rabies are observed, the mice are euthanized, and brain impressions made on 4-well teflon-coated slides (Cel-Line, Erie Scientific, Portsmouth, NH). After overnight fixation in cold acetone, the samples are subjected to typing via the indirect fluorescent antibody test, using a panel of anti-nucleocapsid monoclonal antibodies (N-MABs) of the Centers for Disease Control and Prevention (CDC, Atlanta, GA, USA), and N-MAb 422-5 of the Wistar Institute (Philadelphia, PA, USA), as described by Smith et al., *Ad. Virus Res.*, 1989; 36:215-253.

[00149] The CDC N-MAb panel was developed primarily for discrimination of RABV antigenic variants (Smith, 1989), such that phylogroup I lyssaviruses represent greater variety of antigenic patterns than phylogroup II lyssaviruses (Table 2). Representatives of all four phylogenetically distant lineages of LBV demonstrate the same pattern (except isolate KE576, which was different by one N-MAb). MOKV is distinguished from LBV in reaction with one N-MAb only. Patterns of SHIBV are different from those of LBV by two N-MAbs (C2 and C4), and from those of MOKV by one N-MAb (C4). The positive reaction of SHIBV with the C2 MAb is similar to that of MOKV, WCBV, and many representatives of RABV (Smith, 1989). In addition, as with other African non-RABV lyssaviruses, SHIBV reacts with the Wistar N-MAb 422-5, in contrast to all other lyssaviruses.

[00150] Table 2:

Antigenic patterns of SHIBV compared to other lyssaviruses by a panel of N-MAbs	
Isolate	N-MAbs
SHIBV	C1 (3-1) +
LBV (LBVNIG1956)	C2 (8-2) +
LBV (LBVSA1982)	C3 (11-1) +
LBV (LBVAFR1999)	C4 (15-2) +
LBV (KE131)	C5 (22-3) +
LBV (KE576)	C6 (23-4) +
MOKV	C7 (24-1) +
RABV (CVS)	C8 (24-10) +
RABV (Fox, Europe) ¹⁾	C9 (52-1) +
EBLV-1 ¹⁾	C10 (52-2) +
EBLV-2 ¹⁾	C11 (61-1) +
DUVV ¹⁾	C12 (62-4) +
ARAV ²⁾	C13 (71-2) +
KHUV ²⁾	C14 (15-2) +
IRKV ²⁾	C15 (97-3) +
WCBV ²⁾	C16 (97-11) +
	C17 (141-1) +
	C18 (143-1) +
	C19 (146-3) +
	C20 (164-2) +
	422-5 +

¹⁾ Patterns obtained from Smith (1989). ²⁾ Patterns obtained from Botvinkin et al. (2003).

[00151] Methods involving conventional biological techniques are described herein. Such techniques are generally known in the art and are described in detail in methodology treatises such as Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, ed. Sambrook et al., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989; and Current Protocols in Molecular Biology, ed. Ausubel et al., Greene Publishing and Wiley-Interscience, New York,

1992 (with periodic updates). Immunological methods (e.g., preparation of antigen-specific antibodies, immunoprecipitation, and immunoblotting) are described, e.g., in Current Protocols in Immunology, ed. Coligan et al., John Wiley & Sons, New York, 1991; and Methods of Immunological Analysis, ed. Masseyeff et al., John Wiley & Sons, New York, 1992.

5 **[00152]** Various modifications of the present invention, in addition to those shown and described herein, will be apparent to those skilled in the art of the above description. Such modifications are also intended to fall within the scope of the appended claims.

10 **[00153]** Patents and publications mentioned in the specification are indicative of the levels of those skilled in the art to which the invention pertains. These patents and publications are incorporated herein by reference to the same extent as if each individual application or publication was specifically and individually incorporated herein by reference for the material for which it is cited as well as all other teaching contained therein.

15 **[00154]** The foregoing description is illustrative of particular embodiments of the invention, but is not meant to be a limitation upon the practice thereof. The following claims, including all equivalents thereof, are intended to define the scope of the invention.

REFERENCE LIST

- Badrane, H., Bahloul, C , Perrin, P., Tordo, N., 2001. Evidence of two lyssavirus phylogroups with distinct pathogenecity and immunogenecity. J. Virol. 75, 3268-3276.
- 5 Bahloul, C , Jacob, Y., Tordo, N., Perrin, P. 1998. DNA-based immunization for exploring the enlargement of immunological cross-reactivity against the lyssaviruses. Vaccine 16, 417-425.
- Belikov, S.I., Leonova, G.N., Kondratov, I.G., Romanova, E.V., Pavlenko, E.V., 2009. Isolation and genetic characterisation of a new lyssavirus strain in the Primorskiy kray. East
10 Siberian J. Infect. Pathol. 16(3), 68-69.
- Benmansour, A., Leblois, H., Coulon, P., Tuffereau, C , Gaudin, Y., Flamand, A., Lafay, F., 1991. Antigenicity of rabies virus glycoprotein. J. Virol. 65(8), 4198-4203.
- 15 Botvinkin, A.D., Poleschuk, E.M., Kuzmin, I.V., Borisova, T.I., Gazaryan, S.V., Yager, P. and Rupprecht, C.E., 2003. Novel lyssavirus isolated from bats Russia. Emerging Infect. Dis. 9, 1623-1625.
- Bourhy, H., Kissi, B. and Tordo, N., 1993. Molecular diversity of the *Lyssavirus* genus.
20 Virology 194, 70-81.
- Biichen-Osmond, C , 2003. Taxonomy and Classification of Viruses. In: Murray, P.R., Baron, E.J., Jorgensen, J. H., Pfaller, M.A., Tenover, R. C. (Eds.). Manual of Clinical Microbiology, 8th Edition. Vol 2. ASM Press, Washington, DC, pp. 1217-1226.
- 25 Dean, D.J., M. K. Ableseth, Atanasiu, P., 1996. The fluorescent antibody test. In: Meslin, F.-X., Kaplan, M.M., Koprowski, H. (Eds.). Laboratory techniques in rabies, 4th ed. WHO, Geneva, Switzerland, pp. 88-93.
- 30 Delmas, O., Holmes, E.C., Talbi, C , Larrous, F., Dacheux, L., Bouchier, C , Bourhy, H., 2008. Genomic diversity and evolution of the lyssaviruses. PloS One. 3(4), e2057.

Dietzschold, B., Wunner, W.H., Wiktor, T.J., Lopes, A.D., Lafon, M., Smith, C.L., Koprowski, H. 1983 Characterization of an antigenic determinant of the glycoprotein that correlates with pathogenicity of rabies virus. *Proc. Natl. Acad. Sci. USA.* 80, 70-74.

5 Dietzschold, B., Rupprecht, C.R., Tollis, M., Lafon, M., Mattei, J., Wiktor, T.J. and Koprowski, H., 1988. Antigenic diversity of the glycoprotein and nucleocapsid proteins of rabies and rabies-related viruses: implications for epidemiology and control of rabies. *Rev. Infect. Dis.* 10 (Suppl. 4), 785-798.

10 Drummond, A.J., Rambaut, A., 2007 BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7, 214.

Faber, M., Faber, M.L., Papaneri, A., Bette, M., Weihe, E., Dietzschold, B., Schnell, M.J., 2005. A single amino acid change in rabies virus glycoprotein increases virus spread and
15 enhances virus pathogenicity. *J. Virol.* 79(22), 14141-14148.

Familusi, J.B., Moore, D.L., 1972. Isolation of a rabies related virus from the cerebrospinal fluid of a child with 'aseptic meningitis'. *African J. Med. Sci.* 3, 93-96.

20 Familusi, J.B., Osunkoya, B.O., Moore, D.L., Kemp, G.E., Fabiyi, A., 1972. A fatal human infection with Mokola virus. *Am. J. Trop. Med. Hyg.* 21, 959-963.

Felsenstein, J. 1993. PHYLIP: Phylogenetic inference package (version 3.5c) Dept. of Genetics, University of Washington, Seattle, WA, USA.

25 Fooks, A.R., Brookes, S.M., Johnson, N., McElhinney, L.M., Hutson, A.M., 2003a. European bat lyssaviruses: an emerging zoonosis. *Epidemiol. Infect.* 131(3), 1029-1039.

Fooks, A.R., McElhinney, L.M., Pounder, D.J., Finnegan, C.J., Mansfield, K., Johnson, N.,
30 Brookes, S.M., Parsons, G., White, K., McIntyre, P.G. and Nathwani, D., 2003b. Case report: isolation of a European bat lyssavirus type 2a from a fatal human case of rabies encephalitis. *J. Med. Virol.* 71, 281-289.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl. Acids. Symp. Ser. 41, 95-98.

Hanlon, C.A., Kuzmin, I.V., Blanton, J.D., Weldon, W.C., Manangan, J.S. and Rupprecht, C.E., 2005. Efficacy of rabies biologics against new lyssaviruses from Eurasia. Virus Res. 111, 44-54.

ICTV Official Taxonomy: Updates since the 8th Report. Vertebrate. <http://talk.ictvonline.Org/media/p/1208.aspx> (accessed 10/14/2009).

Jallet, C, Jacob, Y., Bahloul, C, Drings, A., Desmezieres, E., Tordo, N., Perrin, P. 1999. Chimeric lyssavirus glycoproteins with increased immunological potential, J. Virol. 73, 225-233.

Jeanmougin, F., Thompson, J.D., Gouy, M., Higgins, D.G., Gibson, T.J. 1998. Multiple sequence alignment with Clustal X. Trends. Biochem. Sci. 23, 403-405.

Johnson, N., McElhinney, L.M., Smith, J., Lowings, P., Fooks, A.R., 2002. Phylogenetic comparison of the genus *Lyssavirus* using distal coding sequences of the glycoprotein and nucleoprotein genes. Arch. Virol. 147, 2111-2123.

Kissi, B., Tordo, N., Bourhy, H. 1995. Genetic polymorphism in the rabies virus nucleoprotein gene. Virology 209, 526-537.

Koprowski, H., 1996. The mouse inoculation test. In: Meslin, F.-X., Kaplan, M.M., Koprowski, H. (Eds.). Laboratory techniques in rabies, 4th ed. WHO, Geneva, Switzerland, pp. 80-86.

Kumar, S., Tamura, K., Jakobsen, I.B., Nei, M. 2001. MEGA2: molecular evolutionary genetics analysis software. Bioinformatics 17, 1244-1245.

Kuzmin, I.V., Botvinkin, A.D., Poleschuk, E.M., Orciari, L.A., Smith, J.S. and Rupprecht, C.E. (2006). Bat rabies surveillance in the former Soviet Union. Dev. Biol. (Basel). 125, 273-282.

Kuzmin, I.V., Hughes, G.J., Botvinkin, A.D., Orciari, L.A. and Rupprecht, C.E., 2005. Phylogenetic relationships of Irkut and West Caucasian bat viruses within the *Lyssavirus* genus and suggested quantitative criteria based on the N gene sequence for lyssavirus genotype definition. *Virus Res.* 113, 28-43.

Kuzmin, I.V., Niezgoda, M., Franka, R., Agwanda, B., Markotter, W., Beagley, J.C., Urazova, O.Y., Breiman, R.F., Rupprecht, C.E., 2008a. Possible emergence of West Caucasian bat virus in Africa. *Emerging Infect. Dis.* 14(12), 1887-1889.

Kuzmin, I.V., Niezgoda, M., Franka, R., Agwanda, B., Markotter, W., Beagley, J.C., Urazova, O.Y., Breiman, R.F., Rupprecht, C.E., 2008b. Lagos bat virus in Kenya. *J. Clin. Microbiol.* 46(4), 1451-1461.

Kuzmin, I.V., Orciari, L.A., Arai, Y.T., Smith, J.S., Hanlon, C.A., Kameoka, Y. and Rupprecht, C.E., 2003. Bat lyssaviruses (Aravan and Khujand) from Central Asia: phylogenetic relationships according to N, P and G gene sequences. *Virus Res.* 97, 65-79.

Le Mercier, P., Jacob, Y., Tordo, N., 1997. The complete Mokola virus genome sequence: structure of the RNA-dependent RNA polymerase. *J. Gen. Virol.* 78, 1571-1576.

Lentz, T.L., Burrage, T.G., Smith, A.L., Crick, J., Tignor, G.H., 1982. Is the acetylcholine receptor a rabies virus receptor? *Science.* 215(4529), 182-184.

Markotter, W., Kuzmin, I., Rupprecht, C.E., Nel, L.H., 2008a. Phylogeny of Lagos bat virus: challenges for lyssavirus taxonomy. *Virus Res.* 135(1), 10-21.

Markotter, W., Kuzmin, I.V., Rupprecht, C.E., Nel, L.H., 2009. Lagos bat virus virulence in mice inoculated by the peripheral route. *Epidemiol. Infect.* 137(8), 1155-1162.

Markotter, W., Van Eeden, C., Kuzmin, I.V., Rupprecht, C.E., Paweska, J.T., Swanepoel, R., Fooks, A.R., Sabeta, C.T., Cliquet, F., Nel, L.H., 2008b. Epidemiology and pathogenicity of African bat lyssaviruses. *Dev. Biol. (Basel).* 131, 317-325.

Nel, L., Jacobs, J., Jafta, J., Von Teichman, B., Bingham, J. 2000. New Cases of Mokola Virus Infection in South Africa: A Genotypic Comparison of Southern African Virus Isolates. *Virus Genes* 20, 103-106.

5

Nel, L.H., Niezgoda, M., Hanlon, C.A., Morrill, P.A., Yager, P.A., Rupprecht, C.E., 2003. A comparison of DNA vaccines for the rabies-related virus, Mokola. *Vaccine*. 21, 2598-2606.

Pavey, C.R., Burwell, C.J., 1997. The diet of the diadem leaf-nosed bat *Hipposideros diadema*: confirmation of a morphologically-based prediction of carnivory. *J. Zool.* 243(2), 295 - 303.

Poch, O., Blumberg, B. M., Bougueleret, L., Tordo, N., 1990. Sequence comparison of five polymerases (L proteins) of unsegmented negative-strand RNA viruses: theoretical assignment of functional domains. *J. Gen. Virol.* 71, 1153-1162.

15

Prehaud, C , Coulon, P., LaFay, F., Thiers, C , Flamand, A., 1988. Antigenic site II of the rabies virus glycoprotein: structure and role in viral virulence. *J. Virol.* 62(1), 1-7.

Rakotoarivelo, A. A., Ralisata, M., Ravoahangimalala, O.R., Rakotomalala, M.R., Racey, P.A., Jenkins, R.K.B., 2009. The food habits of a Malagasy Giant: *Hipposideros commersoni* (E. Geoffroy, 1813). *African J. Ecol.* 47(3), 283-288.

Rupprecht, C.E., Dietzschold, B., Wunner, W.H. Koprowski, H., 1991. Antigenic relationships of lyssaviruses. In: Baer, G.M. (Ed.) *The natural history of rabies*. 2nd Edition, CRC Press, Boca Raton, pp.69-100.

25

Sabeta, C.T., Markotter, W., Mohale, D.K., Shumba, W., Wandeler, A.I., Nel, L.H., 2007. Mokola virus in domestic mammals, South Africa. *Emerging Inf. Dis.* 13(9), 1371-1373.

30

Shope, R.E. 1982. Rabies-related viruses. *Yale J. Biol. Med.* 55, 271-275.

Shope, R.E., Murphy, F.A., Harrison, A.K., Causey, O.R., Kemp, G.E., Simpson, D.I.H., Moore, D. 1970. Two African viruses serologically and morphologically related to rabies virus. J. Virol. 6, 690-692.

- 5 Smith, J.S., 1989. Rabies virus epitopic variation: use in ecologic studies. Ad. Virus Res. 36, 215-253.

 Swofford, D.L., 2003. PAUP*. Phylogenetic Analysis Using Parsimony (*and other methods) version 4. Sunderland ,MA: Sinauer Associates.

10

 Takayama-Ito, M., Ito, N., Yamada, K., Minamoto, N., Sugiyama, M., 2004. Region at amino acids 164 to 303 of the rabies virus glycoprotein plays an important role in pathogenicity for adult mice. J. Neurovirol. 10(2), 131-135.

- 15 Tignor, G.H., Shope, R.E., Bhatt, P.N., Percy, D.H., 1973. Experimental infection of dogs and monkeys with two rabies serogroup viruses, Lagos bat and Mokola (IbAn 27377): clinical, serologic, virologic and fluorescent-antibody studies. J. Infect. Dis. 128, 471-478.

- Tordo, N., Badrane, H., Bourhy, H., Sacramento, D. 1993. Molecular epidemiology of
20 Lyssaviruses: focus on the glycoprotein and pseudogenes. Onderstepoort J. Vet. Res. 60, 315-323.

- van Thiel, P.P., de Bie, R.M., Eftimov, F., Tepaske, R., Zaaijer, H.L., van Doornum, G.J., Schutten, M., Osterhaus, A.D., Majoie, C.B., Aronica, E., Fehlner-Gardiner, C , Wandeler, A.I.,
25 Kager, P.A., 2009. Fatal human rabies due to Duvenhage Virus from a bat in Kenya: failure of treatment with coma-induction, ketamine, and antiviral drugs. PLoS Negl. Trop. Dis. 3(7), e428.

 Vaughan, T., 1977. Foraging behaviour of the giant leaf-nosed bat (*Hipposideros commersoni*). African J. Ecol. 15(4), 237 - 249.

30

 Warrilow, D., 2005. Australian bat lyssavirus: a recently discovered new rhabdovirus. Curr. Top. Microbiol. Immunol. 292, 25-44.

Webster W.A., Casey, G.A., 1996. Virus isolation in neuroblastoma cell culture. In: Meslin, F.-X., Kaplan, M.M., Koprowski, H. (Eds.). Laboratory techniques in rabies, 4th ed. WHO, Geneva, Switzerland, pp. 96-104.

- 5 Wiktor, T.J., Koprowski, H., 1980. Antigenic variants of rabies virus. J. Exp. Med. 152(1), 99-112.

Wu, X., Franka, R., Velasco-Villa, A., Rupprecht, C.E., 2007. Are all lyssavirus genes equal for phylogenetic analyses? Virus Res. 129(2), 91-103.

CLAIMS

1. An isolated nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO: 1, or a complement thereof.

5

2. A portion of the nucleic acid molecule of SEQ ID NO: 1 corresponding to at least one of the lyssavirus structural genes N, P, M, G, or L.

3. An isolated protein with the sequence of SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, a fragment thereof, or derivative thereof.

10

4. An isolated peptide sequence that is SEQ ID NO: 5, a fragment thereof, or a derivative thereof.

5. A process for detecting the presence of a lyssavirus virus or a nucleic acid molecule derived from a lyssavirus in a biological sample comprising:

15

(a) contacting a sample with an agent that selectively binds to the a nucleic acid molecule with a sequence comprising at least a portion of SEQ ID NO: 1; or a protein sequence encoded by at least a portion of SEQ ID NO: 1; and

(b) detecting whether the agent binds to the virus or the nucleic acid molecule derived therefrom in the sample.

20

6. The method of claim 5, wherein the agent is an antibody.

7. The method of claim 6, wherein the agent is a nucleic acid molecule comprising a nucleotide sequence having between 4 and 6600 contiguous nucleotides of the nucleotide sequence of SEQ ID NO: 1, or a complement thereof.

25

8. A vaccine composition for inducing an immunological response against a lyssavirus in a subject, comprising:

a pharmaceutically acceptable carrier admixed with an isolated lyssavirus strain of Shimoni bat virus.

30

9. A vaccine composition for inducing an immunological response against a lyssavirus in a subject, comprising:

a pharmaceutically acceptable carrier admixed with an isolated protein sequence encoded by at least a portion of SEQ ID NO: 1.

5

10. A vaccine composition for inducing an immunological response against a lyssavirus in a subject, comprising:

a pharmaceutically acceptable carrier admixed with an isolated nucleic acid sequence that has the sequence of least a portion of SEQ ID NO: 1, or derivative thereof.

10

11. The composition of claim 9 wherein said isolated protein sequence is SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, or a derivative thereof.

12. The vaccine composition of any of claims 8-11 formulated for parenteral administration to a subject.

15

13. The vaccine composition of any of claims 8-11 formulated for oral administration to a subject.

14. The vaccine composition of any of claims 8-11 further comprising an adjuvant.

20

15. An isolated lyssavirus virus comprising a nucleic acid molecule having a nucleotide sequence of SEQ ID NO: 1 wherein each thymine is replaced with uracil.



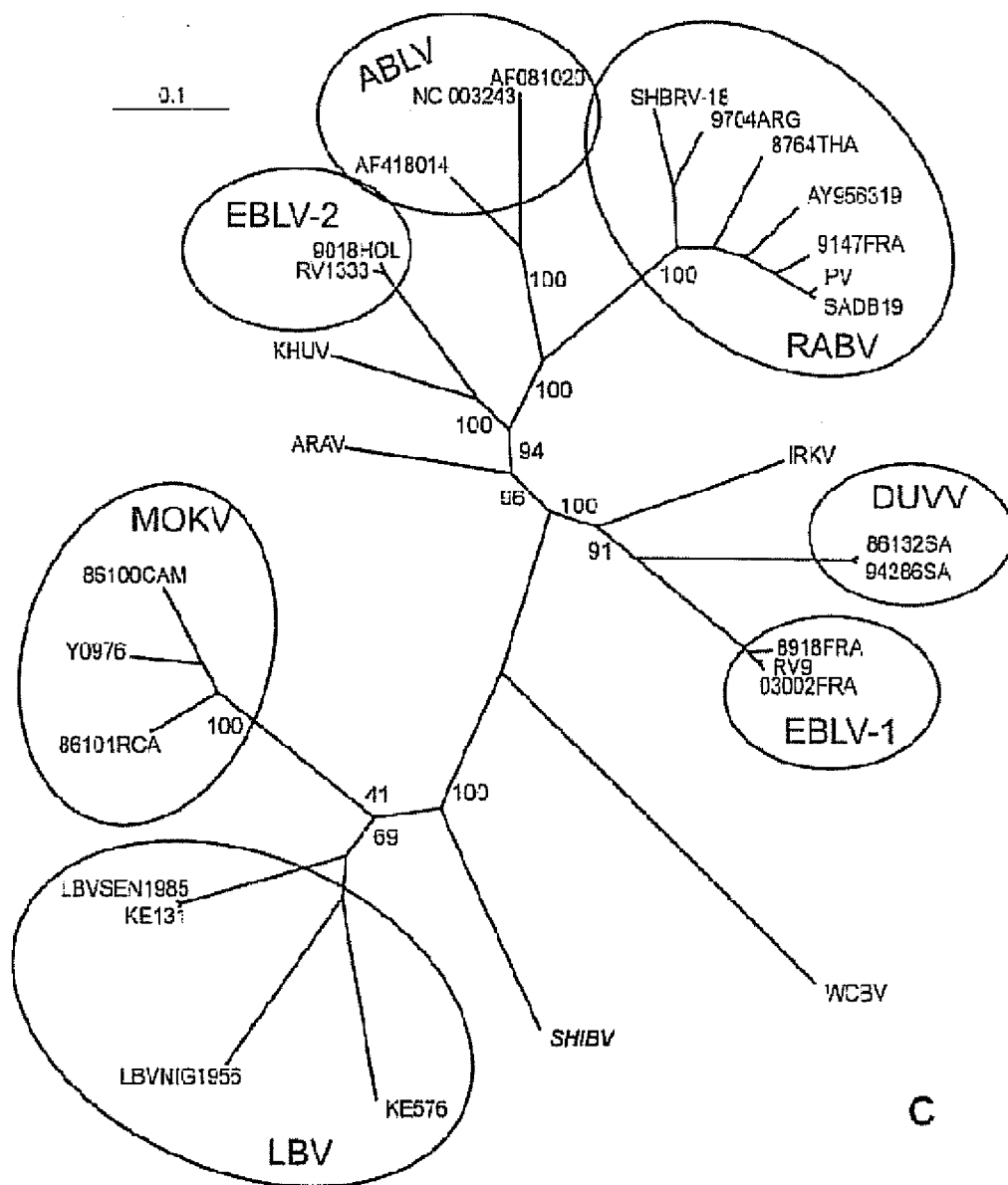


FIG. 1C