

Genetic Detection and Characterization of Lujo Virus, a New Hemorrhagic Fever–Associated Arenavirus from Southern Africa

Thomas Briese¹*, Janusz T. Paweska²*, Laura K. McMullan³, Stephen K. Hutchison⁴, Craig Street¹, Gustavo Palacios¹, Marina L. Khristova⁵, Jacqueline Weyer², Robert Swanepoel², Michael Egholm⁴, Stuart T. Nichol³, W. Ian Lipkin¹*

1 Center for Infection and Immunity, Mailman School of Public Health, Columbia University, New York, New York, United States of America, **2** Special Pathogens Unit, National Institute for Communicable Diseases of the National Health Laboratory Service, Sandringham, South Africa, **3** Special Pathogens Branch, Division of Viral and Rickettsial Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America, **4** 454 Life Sciences, Branford, Connecticut, United States of America, **5** Biotechnology Core Facility Branch, Centers for Disease Control and Prevention, Atlanta, Georgia, United States of America

Abstract

Lujo virus (LUJV), a new member of the family *Arenaviridae* and the first hemorrhagic fever–associated arenavirus from the Old World discovered in three decades, was isolated in South Africa during an outbreak of human disease characterized by nosocomial transmission and an unprecedented high case fatality rate of 80% (4/5 cases). Unbiased pyrosequencing of RNA extracts from serum and tissues of outbreak victims enabled identification and detailed phylogenetic characterization within 72 hours of sample receipt. Full genome analyses of LUJV showed it to be unique and branching off the ancestral node of the Old World arenaviruses. The virus G1 glycoprotein sequence was highly diverse and almost equidistant from that of other Old World and New World arenaviruses, consistent with a potential distinctive receptor tropism. LUJV is a novel, genetically distinct, highly pathogenic arenavirus.

Citation: Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, et al. (2009) Genetic Detection and Characterization of Lujo Virus, a New Hemorrhagic Fever–Associated Arenavirus from Southern Africa. *PLoS Pathog* 4(5): e1000455. doi:10.1371/journal.ppat.1000455

Editor: Michael J. Buchmeier, University of California Irvine, United States of America

Received: February 23, 2009; **Accepted:** April 28, 2009; **Published:** May 29, 2009

This is an open-access article distributed under the terms of the Creative Commons Public Domain declaration which stipulates that, once placed in the public domain, this work may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose.

Funding: This work was supported by Google.org and National Institutes of Health awards AI051292 and AI57158 (Northeast Biodefense Center - Lipkin). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: SKH and ME are employees of 454 Life Sciences, Inc., a Roche Company.

* E-mail: thomas.briese@columbia.edu (TB); wil2001@columbia.edu (WIL)

† These authors contributed equally to this work.

Introduction

Members of the genus *Arenavirus*, comprising currently 22 recognized species (<http://www.ictvonline.org/virusTaxonomy.asp?version=2008>), are divided into two complexes based on serologic, genetic, and geographic relationships [1,2]: the New World (NW) or Tacaribe complex, and the Old World (OW) or Lassa-Lymphocytic choriomeningitis complex that includes the ubiquitous arenavirus type-species *Lymphocytic choriomeningitis virus* (LCMV; [3]). The RNA genome of arenaviruses is bi-segmented, comprising a large (L) and a small (S) segment that each codes for two proteins in ambisense coding strategy [4,5]. Despite this coding strategy, the *Arenaviridae* are classified together with the families *Orthomyxoviridae* and *Bunyaviridae* as segmented single-strand, negative sense RNA viruses.

The South American hemorrhagic fever viruses Junin (JUNV; [6,7]), Machupo (MACV; [8]), Guanarito (GTOV; [9]) and Sabia virus (SABV, [10]), and the African Lassa virus (LASV [11]), are restricted to biosafety level 4 (BSL-4) containment due to their associated aerosol infectivity and rapid onset of severe disease. With the possible exception of NW Tacaribe virus (TCRV; [12]), which has been isolated from bats (*Artibeus* spp.), individual arenavirus species are commonly transmitted by specific rodent species wherein the capacity for persistent infection without overt

disease suggests long evolutionary adaptation between the agent and its host [1,13–16]. Whereas NW arenaviruses are associated with rodents in the *Sigmodontinae* subfamily of the family *Cricetidae*, OW arenaviruses are associated with rodents in the *Murinae* subfamily of the family *Muridae*.

Humans are most frequently infected through contact with infected rodent excreta, commonly via inhalation of dust or aerosolized virus-containing materials, or ingestion of contaminated foods [13]; however, transmission may also occur by inoculation with infected body fluids and tissue transplantation [17–19]. LCMV, which is spread by the ubiquitous *Mus musculus* as host species and hence found world-wide, causes symptoms in humans that range from asymptomatic infection or mild febrile illness to meningitis and encephalitis [13]. LCMV infection is only rarely fatal in immunocompetent adults; however, infection during pregnancy bears serious risks for mother and child and frequently results in congenital abnormalities. The African LASV, which has its reservoir in rodent species of the *Mastomys* genus, causes an estimated 100,000–500,000 human infections per year in West African countries (Figure 1). Although Lassa fever is typically sub-clinical or associated with mild febrile illness, up to 20% of cases may have severe systemic disease culminating in fatal outcome [20,21]. Three other African arenaviruses are not known to cause human disease: Ippy virus (IPPYV; [22,23]), isolated from

Author Summary

In September and October 2008, five cases of undiagnosed hemorrhagic fever, four of them fatal, were recognized in South Africa after air transfer of a critically ill index case from Zambia. Serum and tissue samples from victims were subjected to unbiased pyrosequencing, yielding within 72 hours of sample receipt, multiple discrete sequence fragments that represented approximately 50% of a prototypic arenavirus genome. Thereafter, full genome sequence was generated by PCR amplification of intervening fragments using specific primers complementary to sequence obtained by pyrosequencing and a universal primer targeting the conserved arenaviral termini. Phylogenetic analyses confirmed the presence of a new member of the family *Arenaviridae*, provisionally named Lujo virus (LUJV) in recognition of its geographic origin (Lusaka, Zambia, and Johannesburg, South Africa). Our findings enable the development of specific reagents to further investigate the reservoir, geographic distribution, and unusual pathogenicity of LUJV, and confirm the utility of unbiased high throughput pyrosequencing for pathogen discovery and public health.

Arvicantthis spp. and Mobala virus (MOBV; [24]) isolated from *Praomys* spp. in the Central African Republic (CAR); and Mopeia virus (MOPV) that like LASV is associated with members of the genus *Mastomys*, and was reported from Mozambique [25] and Zimbabwe [26], although antibody studies suggest that MOPV and LASV may also circulate in CAR [27] where the geographies of these viruses appear to overlap (Figure 1). Up to present, there have been no published reports of severe human disease associated with arenaviruses isolated from southern Africa.

In September 2008 an outbreak of unexplained hemorrhagic fever was reported in South Africa [28]. The index patient was airlifted in critical condition from Zambia on September 12 to a clinic in Sandton, South Africa, after infection from an unidentified source. Secondary infections were recognized in a paramedic (case 2) who attended the index case during air transfer from Zambia, in a nurse (case 3) who attended the index case in the intensive care unit in South Africa, and in a member of the hospital staff (case 4) who cleaned the room after the index case died on September 14. One case of tertiary infection was recorded in a nurse (case 5) who attended case 2 after his transfer from Zambia to Sandton on September 26, one day before barrier nursing was implemented. The course of disease in cases 1 through 4 was fatal; case 5 received ribavirin treatment and recovered. A detailed description of clinical and epidemiologic data, as well as immunohistological and PCR analyses that indicated the presence of an arenavirus, are reported in a parallel communication (Paweska et al., *Emerg. Inf. Dis.*, submitted). Here we report detailed genetic analysis of this novel arenavirus.

Results/Discussion

Rapid identification of a novel pathogen through unbiased pyrosequencing

RNA extracts from two post-mortem liver biopsies (cases 2 and 3) and one serum sample (case 2) were independently submitted for unbiased high-throughput pyrosequencing. The libraries yielded between 87,500 and 106,500 sequence reads. Alignment of unique singleton and assembled contiguous sequences to the GenBank database (<http://www.ncbi.nlm.nih.gov/Genbank>) using the Basic Local Alignment Search Tool (blastn and blastx;

[29]) indicated coverage of approximately 5.6 kilobases (kb) of sequence distributed along arenavirus genome scaffolds: 2 kb of S segment sequence in two fragments, and 3.6 kb of L segment sequence in 7 fragments (Figure 2). The majority of arenavirus sequences were obtained from serum rather than tissue, potentially reflecting lower levels of competing cellular RNA in random amplification reactions.

Full genome characterization of a newly identified arenavirus

Sequence gaps between the aligned fragments were rapidly filled by specific PCR amplification with primers designed on the pyrosequence data at both, CU and CDC. Terminal sequences were added by PCR using a universal arenavirus primer, targeting the conserved viral termini (5'-CGC ACM GDG GAT CCT AGG C, modified from [30]) combined with 4 specific primers positioned near the ends of the 2 genome segments. Overlapping primer sets based on the draft genome were synthesized to facilitate sequence validation by conventional dideoxy sequencing. The accumulated data revealed a classical arenavirus genome structure with a bi-segmented genome encoding in an ambisense strategy two open reading frames (ORF) separated by an intergenic stem-loop region on each segment (Figure 2) (GenBank Accession numbers FJ952384 and FJ952385).

Our data represent genome sequences directly obtained from liver biopsy and serum (case 2), and from cell culture isolates obtained from blood at CDC (case 1 and 2), and from liver biopsies at NICD (case 2 and 3). No sequence differences were uncovered between virus detected in primary clinical material and virus isolated in cell culture at the two facilities. In addition, no changes were detected between each of the viruses derived from these first three cases. This lack of sequence variation is consistent with the epidemiologic data, indicating an initial natural exposure of the index case, followed by a chain of nosocomial transmission among subsequent cases.

Lujo virus (LUJV) is a novel arenavirus

Phylogenetic trees constructed from full L or S segment nucleotide sequence show LUJV branching off the root of the OW arenaviruses, and suggest it represents a highly novel genetic lineage, very distinct from previously characterized virus species and clearly separate from the LCMV lineage (Figure 3A and 3B). No evidence of genome segment reassortment is found, given the identical placement of LUJV relative to the other OW arenaviruses based on S and L segment nucleotide sequences. In addition, phylogenetic analysis of each of the individual ORFs reveals similar phylogenetic tree topologies. A phylogenetic tree constructed from deduced L-polymerase amino acid (aa) sequence also shows LUJV near the root of the OW arenaviruses, distinct from characterized species, and separate from the LCMV branch (Figure 3C). A distant relationship to OW arenaviruses may also be inferred from the analysis of Z protein sequence (Figure S1). The NP gene sequence of LUJV differs from other arenaviruses from 36% (IPPYV) to 43% (TAMV) at the nucleotide level, and from 41% (MOBV/LASV) to 55% (TAMV) at the aa level (Table S1). This degree of divergence is considerably higher than both, proposed cut-off values within (<10–12%), or between (>21.5%) OW arenavirus species [31,32], and indicates a unique phylogenetic position for LUJV (Figure 3D). Historically, phylogenetic assignments of arenaviruses have been based on portions of the NP gene [1,33], because this is the region for which most sequences are known. However, as more genomic sequences have become available, analyses of full-length GPC sequence have revealed evidence of possible relationships between OW and NW

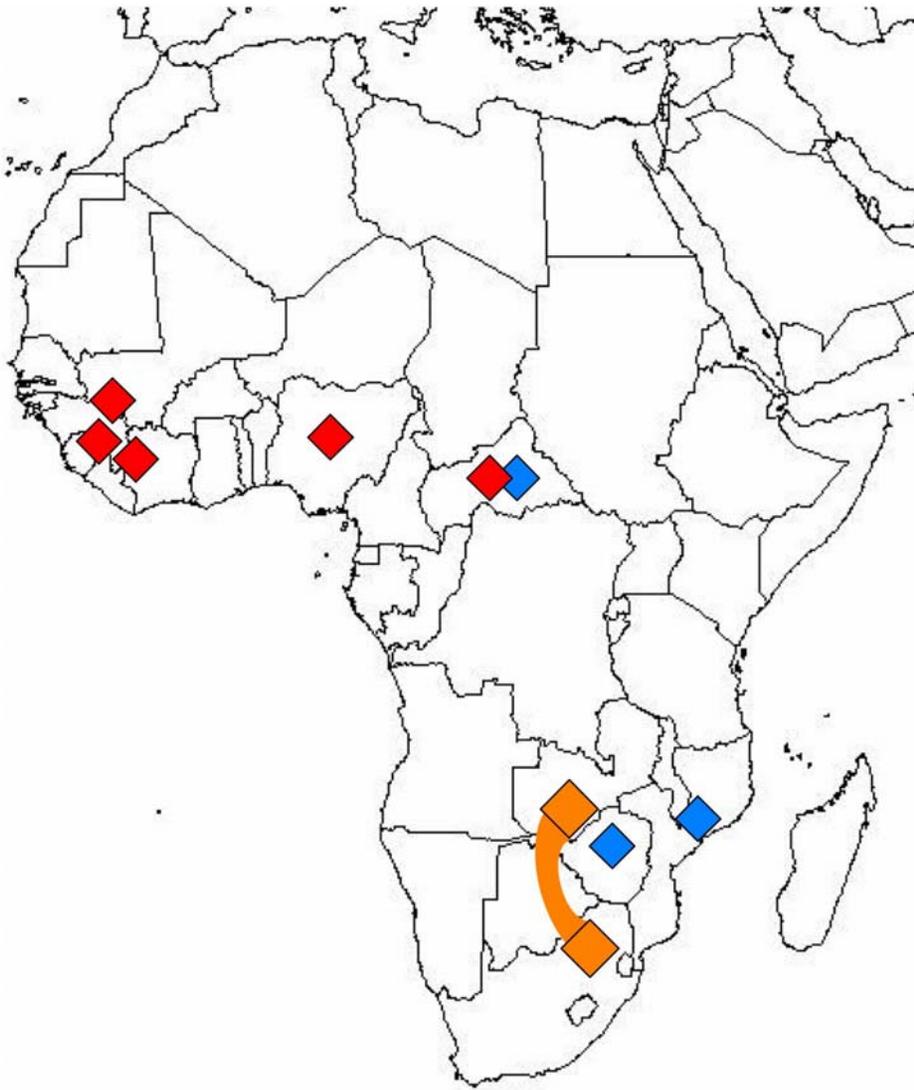


Figure 1. Geographic distribution of African arenaviruses. MOBV, MOPV, and IPPYV (blue) have not been implicated in human disease; LASV (red) can cause hemorrhagic fever. The origin of the LUJV index and secondary and tertiary cases linked in the 2008 outbreak are indicated in gold. doi:10.1371/journal.ppat.1000455.g001

arenaviruses not revealed by NP sequence alone [34]. Because G1 sequences are difficult to align some have pursued phylogenetic analyses by combining the GPC signal peptide and the G2 sequence for phylogenetic analysis [16]. We included in our analysis the chimeric signal/G2 sequence (Figure 3E) as well as the receptor binding G1 portion (Figure 3F); both analyses highlighted the novelty of LUJV, showing an almost similar distance from OW as from NW viruses.

Protein motifs potentially relevant to LUJV biology

Canonical polymerase domains pre-A, A, B, C, D, and E [35–37] are well conserved in the L ORF of LUJV (256 kDa, pI = 6.4; Figure 4). The Z ORF (10.5 kDa, pI = 9.3) contains two late domain motifs like LASV; however, in place of the PTAP motif found in LASV, that mediates recognition of the tumor susceptibility gene 101, Tsg101 [38], involved in vacuolar protein sorting [39,40], LUJV has a unique Y₇₇REL motif that matches the YXXL motif of the retrovirus equine infectious anemia virus

[41], which interacts with the clathrin adaptor protein 2 (AP2) complex [42]. A Tsg101-interacting motif, P₉₀SAP, is found in LUJV in position of the second late domain of LASV, PPPY, which acts as a Nedd4-like ubiquitin ligase recognition motif [43]. The RING motif, containing conserved residue W₄₄ [44], and the conserved myristoylation site G₂ are present [45–47] (Figure 4). The NP of LUJV (63.1 kDa, pI = 9.0) contains described aa motifs that resemble mostly OW arenaviruses [48], including a cytotoxic T-lymphocyte (CTL) epitope reported in LCMV (GVYMGNL; [49]), corresponding to G₁₂₂VYRGNL in LUJV, and a potential antigenic site reported in the N-terminal portion of LASV NP (RKSKRND; [50]), corresponding to R₅₅KDKRND in LUJV (Figure 4).

The GPC precursor (52.3 kDa, pI = 9.0) is cotranslationally cleaved into the long, stable signal peptide and the mature glycoproteins G1 and G2 [51–54]. Based on analogy to LASV [55] and LCMV [56], signalase would be predicted to cleave between D₅₈ and S₅₉ in LUJV. However, aspartate and arginine

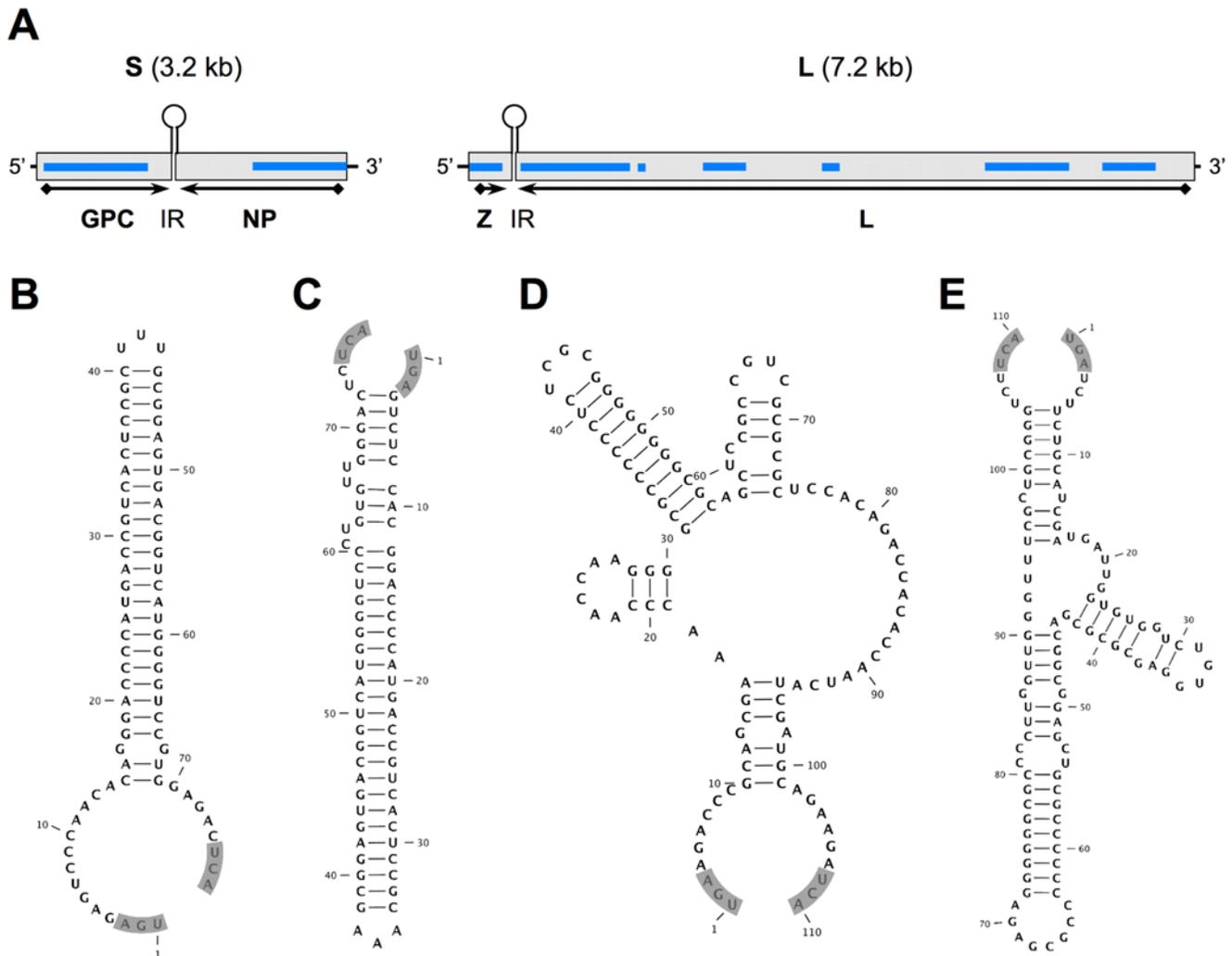


Figure 2. LUJV genome organization and potential secondary structure of intergenic regions. Open reading frames (ORF) for the glycoprotein precursor GPC, the nucleoprotein NP, the matrix protein analog Z, and the polymerase L, and their orientation are indicated (A); blue bars represent sequences obtained by pyrosequencing from clinical samples. Secondary structure predictions of intergenic regions (IR) for S (B, C) and L segment sequence (D, E) in genomic (B, D) and antigenomic orientation (C, E) were analyzed by mfold; shading indicates the respective termination codon (opal, position 1), and its reverse-complement, respectively.
doi:10.1371/journal.ppat.1000455.g002

residues in the -1 and -3 positions, respectively, violate the $(-3, -1)$ -rule [57]; thus, cleavage may occur between S_{59} and S_{60} as predicted by the SignalP algorithm. The putative 59 aa signal peptide of LUJV displays a conserved G_2 , implicated in myristoylation in JUNV [58], however, it is followed in LUJV by a non-standard valine residue in position +4, resembling non-standard glycine residues found in Oliveros virus (OLVV [59]) and Latino virus (LATV; <http://www2.ncid.cdc.gov/arbocat/catalog-listing.asp?VirusID=263&SI=1>). Conservation is also observed for aa residues P_{12} (except Amapari virus; AMAV [60]), E_{17} [61] (except Pirital virus; PIRV [62]), and N_{20} in hydrophobic domain 1, as well as $I_{32}KGVFNLYK_{40}SG$, identified as a CTL epitope in LCMV WE ($I_{32}KAVYNFATCG$; [63]) (Figure 4).

Analogous to other arenaviruses, SKI-1/S1P cleavage C-terminal of $RKLM_{221}$ is predicted to separate mature G1 (162 aa, 18.9 kDa, $pI=6.4$) from G2 (233 aa, 26.8 kDa, $pI=9.5$) [52,53,64]. G2 appears overall well conserved, including the strictly conserved cysteine residues: 6 in the luminal domain, and 3 in the cytoplasmic tail that are included in a conserved zinc finger

motif reported in JUNV [65] (Figure 4). G2 contains 6 potential glycosylation sites, including 2 strictly conserved sites, 2 semi-conserved sites N_{335} (absent in LCMVs and Dandenong virus; DANV [19]) and N_{352} (absent in LATV), and 2 unique sites in the predicted cytoplasmic tail (Figure 4). G1 is poorly conserved among arenaviruses [16], and G1 of LUJV is no exception, being highly divergent from the G1 of the other arenaviruses, and shorter than that of other arenaviruses. LUJV G1 contains 6 potential glycosylation sites in positions comparable to other arenaviruses, including a conserved site $N_{93}HS$ (Figure 4), which is shifted by one aa in a motif that otherwise aligns well with OW arenaviruses and NW arenavirus clade A and C viruses. There is no discernable homology to other arenavirus G1 sequences that would point to usage of one of the two identified arenavirus receptors; Alpha-dystroglycan (α -DG) [66] that binds OW arenaviruses LASV and LCMV, and NW clade C viruses OLVV and LATV [67], or transferrin receptor 1 (TfR1) that binds pathogenic NW arenaviruses JUNV, MACV, GTOV, and SABV [68] (Figure S2).

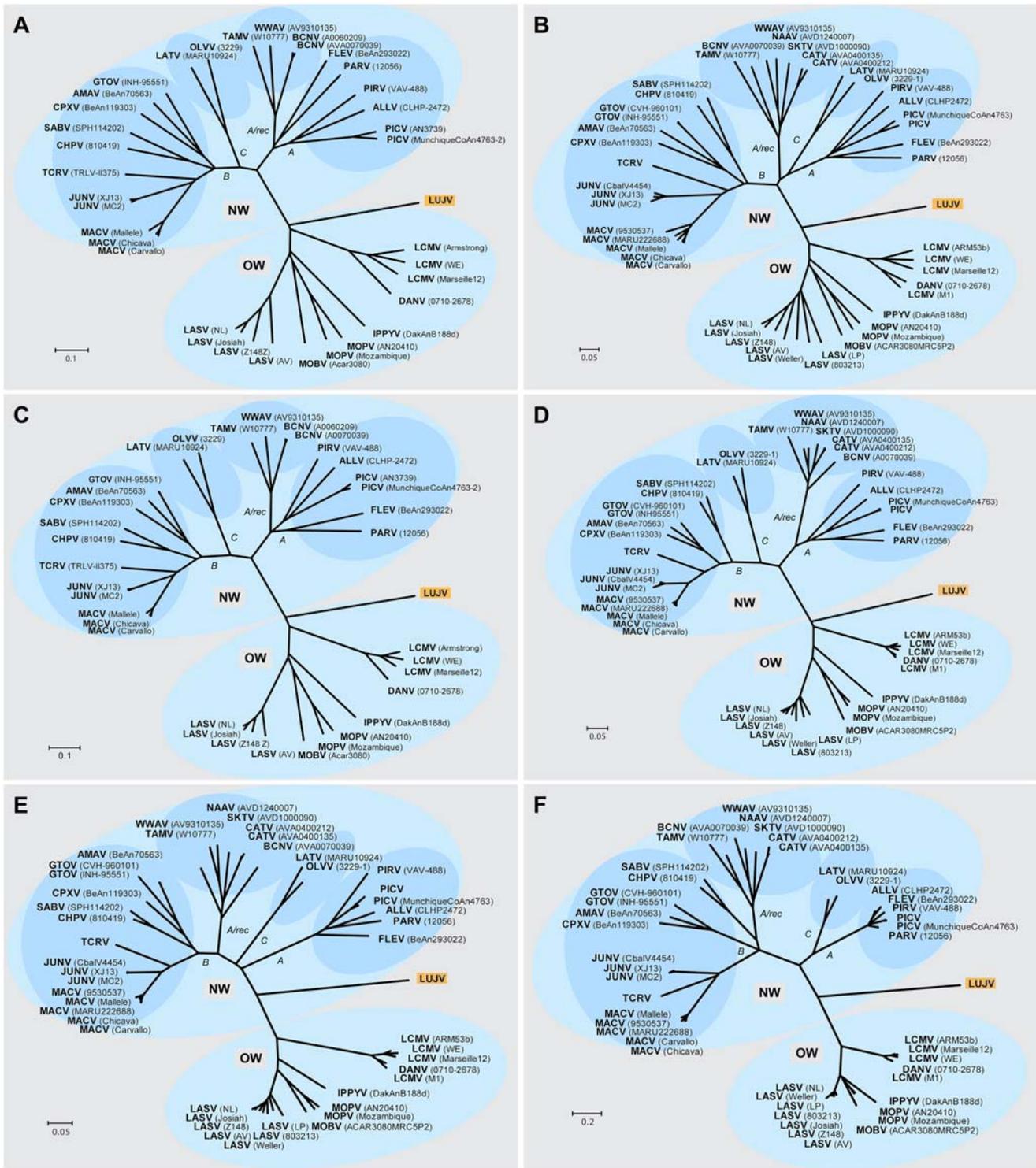


Figure 3. Phylogenetic analyses of LUJV. Phylogenetic relationships of LUJV were inferred based on full L (A) and S segment nucleotide sequence (B), as well as on deduced amino acid sequences of L (C), NP (D), Signal/G2 (E) and G1 (F) ORF's. Phylogenies were reconstructed by neighbor-joining analysis applying a Jukes-Cantor model; the scale bar indicates substitutions per site; robust bootstrap support for the positioning of LUJV was obtained in all cases (>98% of 1000 pseudoreplicates). GenBank Accession numbers for reference sequences are: ALLV CLHP2472 (AY216502, AY012687); AMAV BeAn70563 (AF512834); BCNV AVA0070039 (AY924390, AY922491), A0060209 (AY216503); CATV AVA0400135 (DQ865244), AVA0400212 (DQ865245); CHPV 810419 (EU, 260464, EU260463); CPXV BeAn119303 (AY216519, AF512832); DANV 0710-2678 (EU136039, EU136038); FLEV BeAn293022 (EU627611, AF512831); GTOV INH-95551 (AY358024, AF485258), CVH-960101 (AY497548); IPPYV DakAnB188d (DQ328878, DQ328877); JUNV MC2 (AY216507, D10072), XJ13 (AY358022, AY358023), CbalV4454 (DQ272266); LASV LP (AF181853), 803213 (AF181854), Weller (AY628206), AV (AY179171, AF246121), Z148 (AY628204, AY628205), Josiah (U73034, J043200), NL (AY179172, AY179173); LATV MARU10924 (EU627612, AF485259); LCMV Armstrong (AY847351), ARM53b (M20869), WE (AF004519, M22138), Marseille12 (DQ286932, DQ286931), M1 (AB261991); MACV Carvalho (AY619642, AY619643), Chicava (AY624354, AY624355), Mallele (AY619644, AY619645), MARU22688

(AY922407), 9530537 (AY571959); MOBV ACAR3080MRC5P2 (DQ328876, AY342390); MOPV AN20410 (AY772169, AY772170), Mozambique (DQ328875, DQ328874); NAAV AVD1240007 (EU123329); OLVV 3229-1 (AY216514, U34248); PARV 12056 (EU627613, AF485261); PICV (K02734), MunchiqueCoAn4763 (EF529745, EF529744), AN3739 (AF427517); PIRV VAV-488 (AY216505, AF277659); SABV SPH114202 (AY358026, U41071); SKTV AVD1000090 (EU123328); TAMV W10777 (EU627614, AF512828); TCRV (J04340, M20304); WWAV AV9310135 (AY924395, AF228063). doi:10.1371/journal.ppat.1000455.g003

In summary, our analysis of the LUJV genome shows a novel virus that is only distantly related to known arenaviruses. Sequence divergence is evident across the whole genome, but is most pronounced in the G1 protein encoded by the S segment, a region implicated in receptor interactions. Reassortment of S and L segments leading to changes in pathogenicity has been described in cultured cells infected with different LCMV strains [69], and between pathogenic LASV and nonpathogenic MOPV [70]. We find no evidence to support reassortment of the LUJV L or S genome segment (Figure 3A and 3B). Recombination of glycoprotein sequence has been recognized in NW arenaviruses [14,16,33,34,71–73], resulting in the division of the complex into four sublineages: lineages A, B, C, and an A/recombinant lineage that forms a branch of lineage A when NP and L sequence is considered (see Figure 3C and 3D), but forms an independent branch in between lineages B and C when glycoprotein sequence is considered (see Figure 3D). While recombination cannot be excluded in case of LUJV, our review of existing databases reveals no candidate donor for the divergent GPC sequence. To our knowledge is LUJV the first hemorrhagic fever-associated arenavirus from Africa identified in the past 3 decades. It is also the first such virus originating south of the equator (Figure 1). The International Committee on the Taxonomy of Viruses (ICTV) defines species within the *Arenavirus* genus based on association with a specific host, geographic distribution, potential to cause

human disease, antigenic cross reactivity, and protein sequence similarity to other species. By these criteria, given the novelty of its presence in southern Africa, capacity to cause hemorrhagic fever, and its genetic distinction, LUJV appears to be a new species.

Materials and Methods

Sequencing

Clinical specimens were inactivated in TRIzol (liver tissue, 100 mg) or TRIzol LS (serum, 250 µl) reagent (Invitrogen, Carlsbad, CA, USA) prior to removal from BSL-4 containment. Total RNA extracts were treated with DNase I (DNA-free, Ambion, Austin, TX, USA) and cDNA generated by using the Superscript II system (Invitrogen) and 100–500 ng RNA for reverse transcription primed with random octamers that were linked to an arbitrary, defined 17-mer primer sequence [74]. The resulting cDNA was treated with RNase H and then randomly amplified by the polymerase chain reaction (PCR; [75]); applying a 9:1 mixture of a primer corresponding to the defined 17-mer sequence, and the random octamer-linked 17-mer primer, respectively [74]. Products >70 base pairs (bp) were selected by column purification (MinElute, Qiagen, Hilden, Germany) and ligated to specific linkers for sequencing on the 454 Genome Sequencer FLX (454 Life Sciences, Branford, CT, USA) without fragmentation of the cDNA [19,76,77]. Removal of primer sequences, redundancy filtering,

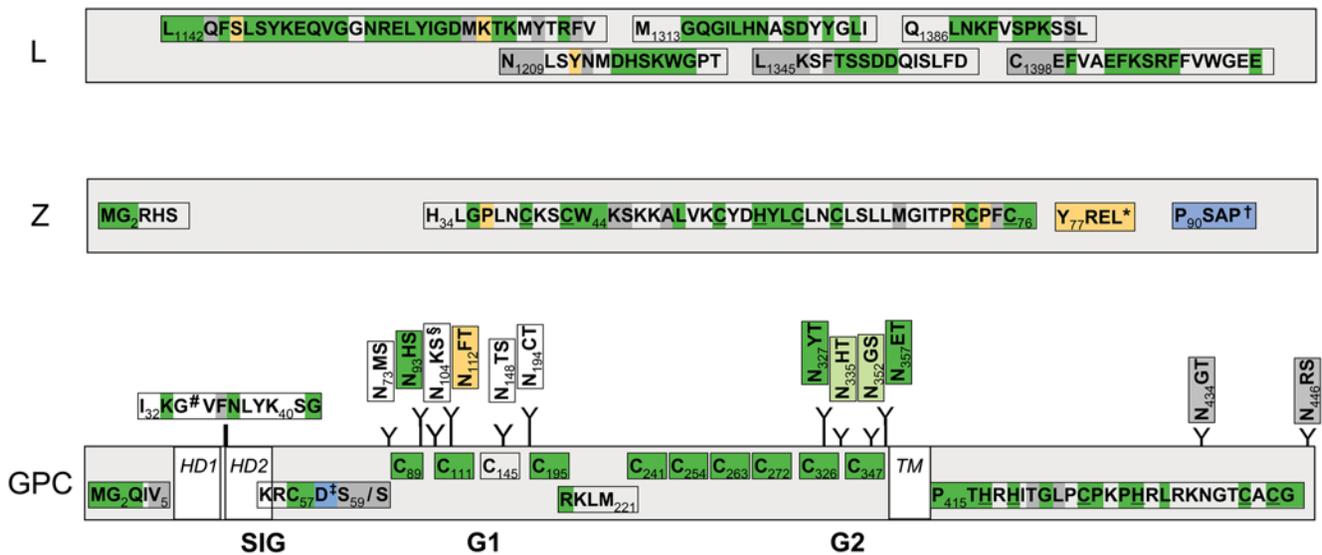


Figure 4. Schematic of conserved protein motifs. Conservation of LUJV amino acid motifs with respect to all other (green highlight), to OW (yellow highlight), or to NW (blue highlight) arenaviruses is indicated; grey highlight indicates features unique to LUJV. Polymerase motifs pre-A (L₁₁₄₂), A (N₁₂₀₉), B (M₁₃₁₃), C (L₁₃₄₅), D (Q₁₃₈₆), and E (C₁₃₉₈) are indicated for the L ORF; potential myristoylation site G₂, the RING motif H₃₄/C₇₆, and potential late domains YXXL an PSAP are indicated for the Z ORF; and myristoylation site G₂, posttranslational processing sites for signalase (S₅₉/S₆₀) and S1P cleavage (RKL M₂₂₁), CTL epitope (I₃₂), zinc finger motif P₄₁₅/G₄₄₀, as well as conserved cysteine residues and glycosylations sites (Y) are indicated for GPC. * late domain absent in NW viruses and DANV; † PSAP or PTAP in NW viruses, except in PIRV and TCRV (OW viruses: PPPY); # G in all viruses except LCMV (= A); ‡ D in NW clade A only; § conserved with respect to OW, and NW clade A and C; HD, hydrophobic domain; TM, transmembrane anchor.

doi:10.1371/journal.ppat.1000455.g004

and sequence assembly were performed with software programs accessible through the analysis applications at the GreenePortal website (<http://156.145.84.111/Tools>).

Conventional PCRs at CU were performed with HotStar polymerase (Qiagen) according to manufacturer's protocols on PTC-200 thermocyclers (Bio-Rad, Hercules, CA, USA): an enzyme activation step of 5 min at 95°C was followed by 45 cycles of denaturation at 95°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 1 to 3 min depending on the expected amplicon size. A two-step RT-PCR protocol was also followed at CDC using Invitrogen's ThermoScript RT at 60 degrees for 30 min followed by RNase H treatment for 20 min. cDNA was amplified using Phusion enzyme with GC Buffer (Finnzymes, Espoo, Finland) and 3% DMSO with an activation step at 98°C for 30 sec, followed by the cycling conditions of 98°C for 10 sec, 58°C for 20 sec, and 72°C for 1 min for 35 cycles and a 5 min extension at 72°C. Specific primer sequences are available upon request. Amplification products were run on 1% agarose gels, purified (MinElute, Qiagen), and directly sequenced in both directions with ABI PRISM Big Dye Terminator 1.1 Cycle Sequencing kits on ABI PRISM 3700 DNA Analyzers (Perkin-Elmer Applied Biosystems, Foster City, CA).

Sequence analyses

Programs of the Wisconsin GCG Package (Accelrys, San Diego, CA, USA) were used for sequence assembly and analysis; percent sequence difference was calculated based on Needleman-Wunsch alignments (gap open/extension penalties 15/6.6 for nucleotide and 10/0.1 for aa alignments; EMBOSS [78]), using a Perl script to iterate the process for all versus all comparison. Secondary RNA structure predictions were performed with the web-based version of mfold (<http://mfold.bioinfo.rpi.edu>); data were exported as .ct files and layout and annotation was done with CLC RNA Workbench (CLC bio, Århus, Denmark). Protein topology and targeting predictions were generated by employing SignalP, and NetNGlyc, TMHMM (<http://www.cbs.dtu.dk/services>), the web-based version of TopPred (<http://mobyle.pasteur.fr/cgi-bin/portal.py?form=toppred>), and Phobius (<http://phobius.sbc.su.se/>). Phylogenetic analyses were performed using MEGA software [79].

References

- Bowen MD, Peters CJ, Nichol ST (1997) Phylogenetic analysis of the Arenaviridae: patterns of virus evolution and evidence for cospeciation between arenaviruses and their rodent hosts. *Mol Phylogenet Evol* 8: 301–316.
- Moncayo AC, Hice CL, Watts DM, Travassos de Rosa AP, Guzman H, et al. (2001) Allpahuayo virus: a newly recognized arenavirus (arenaviridae) from arboreal rice rats (*Oecomys bicolor* and *Oecomys paricola*) in northeastern Peru. *Virology* 284: 277–286.
- Armstrong C, Lillie RD (1934) Experimental lymphocytic choriomeningitis of monkeys and mice produced by a virus encountered in studies of the 1933 St. Louis encephalitis epidemic. *Public Health Rep* 49: 1019–1027.
- Auperin DD, Romanowski V, Galinski M, Bishop DH (1984) Sequencing studies of pichinde arenavirus S RNA indicate a novel coding strategy, an ambisense viral S RNA. *J Virol* 52: 897–904.
- Salvato MS, Shimomaye EM (1989) The completed sequence of lymphocytic choriomeningitis virus reveals a unique RNA structure and a gene for a zinc finger protein. *Virology* 173: 1–10.
- Parodi AS, Greenway DJ, Rugiero HR, Frigerio M, De La Barrera JM, et al. (1958) [Concerning the epidemic outbreak in Junin.]. *Dia Med* 30: 2300–2301.
- Pirosky I, Zuccarini J, Molinelli EA, Di Pietro A, Barrera Oro JG, et al. (1959) Viremia hemorrhagica del Noroeste Bonaerense. *Orientacion Medica* 8: 303–311.
- Johnson KM, Wiebenga NH, Mackenzie RB, Kuns ML, Tauraso NM, et al. (1965) Virus Isolations from Human Cases of Hemorrhagic Fever in Bolivia. *Proc Soc Exp Biol Med* 118: 113–118.
- Salas R, de Manzione N, Tesh RB, Rico-Hesse R, Shope RE, et al. (1991) Venezuelan haemorrhagic fever. *Lancet* 338: 1033–1036.
- Coimbra TLM, Nassar ES, Burattini MN, de Souza LTM, Ferreira IB, et al. (1994) New arenavirus isolated in Brazil. *Lancet* 343: 391–392.

Supporting Information

Figure S1 Phylogenetic tree based on deduced Z amino acid sequence. In contrast to phylogenetic trees obtained with the other ORFs (Figure 2), poor bootstrap support (43% of 1,000 pseudoreplicates) for the branching of LUJV off the LCMV clade was obtained with Z ORF sequence. For GenBank accession numbers see Figure 2.

Found at: doi:10.1371/journal.ppat.1000455.s001 (0.44 MB TIF)

Figure S2 Pairwise sliding-window distance analysis of GPC sequence. LUJV and members of the OW (LASV, MOPV, IPPYV, LCMV, DANV) and NW (GTOV, CPXV, BNCV, PIRV, OLVV, SABV, MACV) arenavirus complex were compared using LASV NL (A) or GTOV CVH (B) as query (10 aa step; 80 aa window).

Found at: doi:10.1371/journal.ppat.1000455.s002 (4.21 MB TIF)

Table S1 Pairwise nucleotide and amino acid differences between LUJV and other OW and NW arenaviruses. * NAAV, North American arenavirus. † Values <30% (amino acid) or <33% (nucleotide) are highlighted in green.

Found at: doi:10.1371/journal.ppat.1000455.s003 (0.20 MB DOC)

Acknowledgments

We are grateful to the Outbreak Control and Investigation Teams in South Africa and Zambia, and to Cathy Roth and the WHO GOARN network for their help and support. We thank Robert Serge, Jeffrey Hui, Alla Tashmukhamedova, and Katrina Ciraldo for technical assistance, and Phenix-Lan Quan for data analysis and critical comments.

Author Contributions

Conceived and designed the experiments: TB WIL. Performed the experiments: TB JTP LKM SKH GP MLK JW. Analyzed the data: TB LKM SKH CS GP MLK ME STN WIL. Contributed reagents/materials/analysis tools: JTP CS JW BS ME. Wrote the paper: TB JTP BS STN WIL.

- Buckley SM, Casals J (1970) Lassa fever, a new virus disease of man from West Africa. 3. Isolation and characterization of the virus. *Am J Trop Med Hyg* 19: 680–691.
- Downs WG, Anderson CR, Spence L, Aitken THG, Greenhall AH (1963) Tacaribe virus, a new agent isolated from Artibeus bats and mosquitoes in Trinidad, West Indies. *Am J Trop Med Hyg* 12: 640–646.
- Buchmeier MJ, de la Torre JC, Peters CJ (2007) Arenaviridae: the viruses and their replication. In: Knipe DM, Howley PM, eds. *Fields Virology*. Philadelphia, PA, USA: Wolter Kluwer Lippincott Williams & Wilkins. pp 1791–1827.
- Fulhorst CF, Bowen MD, Ksiazek TG, Rollin PE, Nichol ST, et al. (1996) Isolation and characterization of Whitewater Arroyo virus, a novel North American arenavirus. *Virology* 224: 114–120.
- Hugot JP, Gonzalez JP, Denys C (2001) Evolution of the Old World Arenaviridae and their rodent hosts: generalized host-transfer or association by descent? *Infect Genet Evol* 1: 13–20.
- Charrel RN, de Lamballerie X, Emonet S (2008) Phylogeny of the genus Arenavirus. *Curr Opin Microbiol* 11: 362–368.
- Fischer SA, Graham MB, Kuehnert MJ, Kotton CN, Srinivasan A, et al. (2006) Transmission of lymphocytic choriomeningitis virus by organ transplantation. *N Engl J Med* 354: 2235–2249.
- Amman BR, Pavlin BI, Albarino CG, Comer JA, Erickson BR, et al. (2007) Pet rodents and fatal lymphocytic choriomeningitis in transplant patients. *Emerg Infect Dis* 13: 719–725.
- Palacios G, Druce J, Du L, Tran T, Birch C, et al. (2008) A new arenavirus in a cluster of fatal transplant-associated diseases. *N Engl J Med* 358: 991–998.
- Ogbu O, Ajuluchukwu E, Uneke CJ (2007) Lassa fever in West African sub-region: an overview. *J Vector Borne Dis* 44: 1–11.

21. Khan SH, Goba A, Chu M, Roth C, Healing T, et al. (2008) New opportunities for field research on the pathogenesis and treatment of Lassa fever. *Antiviral Res* 78: 103–115.
22. Swanepoel R, Leman PA, Shepherd AJ, Shepherd SP, Kiley MP, et al. (1985) Identification of Ippy as a Lassa-fever-related virus. *Lancet* 1: 639.
23. Meunier DY, McCormick JB, Georges AJ, Georges MC, Gonzalez JP (1985) Comparison of Lassa, Mobala, and Ippy virus reactions by immunofluorescence test. *Lancet* 1: 873–874.
24. Gonzalez JP, McCormick JB, Saluzzo JF, Herve JP, Georges AJ, et al. (1983) An arenavirus isolated from wild-caught rodents (*Pramys* species) in the Central African Republic. *Intervirology* 19: 105–112.
25. Wulff H, McIntosh BM, Hamner DB, Johnson KM (1977) Isolation of an arenavirus closely related to Lassa virus from *Mastomys natalensis* in south-east Africa. *Bull World Health Organ* 55: 441–444.
26. Johnson KM, Taylor P, Elliott LH, Tomori O (1981) Recovery of a Lassa-related arenavirus in Zimbabwe. *Am J Trop Med Hyg* 30: 1291–1293.
27. Georges AJ, Gonzalez JP, Abdul-Wahid S, Saluzzo JF, Meunier DM, et al. (1985) Antibodies to Lassa and Lassa-like viruses in man and mammals in the Central African Republic. *Trans R Soc Trop Med Hyg* 79: 78–79.
28. National Institute for Communicable Diseases (2008) Arenavirus outbreak, South Africa. *Communicable Diseases Communique* 7: 1–3. <http://www.nicd.ac.za>.
29. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215: 403–410.
30. Clegg JC, Wilson SM, Oram JD (1991) Nucleotide sequence of the S RNA of Lassa virus (Nigerian strain) and comparative analysis of arenavirus gene products. *Virus Res* 18: 151–164.
31. Bowen MD, Rollin PE, Ksiazek TG, Hustad HL, Bausch DG, et al. (2000) Genetic diversity among Lassa virus strains. *J Virol* 74: 6992–7004.
32. Emonet S, Lemasson JJ, Gonzalez JP, de Lamballerie X, Charrel RN (2006) Phylogeny and evolution of old world arenaviruses. *Virology* 350: 251–257.
33. Bowen MD, Peters CJ, Nichol ST (1996) The phylogeny of New World (Tacaribe complex) arenaviruses. *Virology* 219: 285–290.
34. Albarino CG, Posik DM, Ghiringhelli PD, Lozano ME, Romanowski V (1998) Arenavirus phylogeny: a new insight. *Virus Genes* 16: 39–46.
35. Poch O, Sauvaget I, Delarue M, Tordo N (1989) Identification of four conserved motifs among the RNA-dependent polymerase encoding elements. *EMBO J* 8: 3867–3874.
36. Delarue M, Poch O, Tordo N, Moras D, Argos P (1990) An attempt to unify the structure of polymerases. *Protein Eng* 3: 461–467.
37. Müller R, Poch O, Delarue M, Bishop DH, Bouloy M (1994) Rift Valley fever virus L segment: correction of the sequence and possible functional role of newly identified regions conserved in RNA-dependent polymerases. *J Gen Virol* 75(Pt 6): 1345–1352.
38. Perez M, Craven RC, de la Torre JC (2003) The small RING finger protein Z drives arenavirus budding: implications for antiviral strategies. *Proc Natl Acad Sci U S A* 100: 12978–12983.
39. Garrus JE, von Schwedler UK, Pornillos OW, Morham SG, Zavitz KH, et al. (2001) Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* 107: 55–65.
40. VerPlank L, Bouamr F, LaGrassa TJ, Agresta B, Kikonyogo A, et al. (2001) Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1 Pr55(Gag). *Proc Natl Acad Sci U S A* 98: 7724–7729.
41. Puffer BA, Parent LJ, Wills JW, Montelaro RC (1997) Equine infectious anemia virus utilizes a YXXL motif within the late assembly domain of the Gag p9 protein. *J Virol* 71: 6541–6546.
42. Puffer BA, Watkins SC, Montelaro RC (1998) Equine infectious anemia virus Gag polyprotein late domain specifically recruits cellular AP-2 adapter protein complexes during virion assembly. *J Virol* 72: 10218–10221.
43. Staub O, Dho S, Henry P, Correa J, Ishikawa T, et al. (1996) WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na⁺ channel deleted in Liddle's syndrome. *EMBO J* 15: 2371–2380.
44. Joazeiro CA, Wing SS, Huang H, Levenson JD, Hunter T, et al. (1999) The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science* 286: 309–312.
45. Perez M, Greenwald DL, de la Torre JC (2004) Myristoylation of the RING finger Z protein is essential for arenavirus budding. *J Virol* 78: 11443–11448.
46. Strecker T, Maisa A, Daffis S, Eichler R, Lenz O, et al. (2006) The role of myristoylation in the membrane association of the Lassa virus matrix protein Z. *Virol J* 3: 93.
47. Capul AA, Perez M, Burke E, Kunz S, Buchmeier MJ, et al. (2007) Arenavirus Z-glycoprotein association requires Z myristoylation but not functional RING or late domains. *J Virol* 81: 9451–9460.
48. Gonzalez JP, Bowen MD, Nichol ST, Rico-Hesse R (1996) Genetic characterization and phylogeny of Sabia virus, an emergent pathogen in Brazil. *Virology* 221: 318–324.
49. Whitton JL, Tishon A, Lewicki H, Gebhard J, Cook T, et al. (1989) Molecular analyses of a five-amino-acid cytotoxic T-lymphocyte (CTL) epitope: an immunodominant region which induces nonreciprocal CTL cross-reactivity. *J Virol* 63: 4303–4310.
50. Gonzalez JP, Sanchez A, Rico-Hesse R (1995) Molecular phylogeny of Guanarito virus, an emerging arenavirus affecting humans. *Am J Trop Med Hyg* 53: 1–6.
51. Lenz O, ter Meulen J, Klenk HD, Seidah NG, Garten W (2001) The Lassa virus glycoprotein precursor GP-C is proteolytically processed by subtilase SKI-1/S1P. *Proc Natl Acad Sci U S A* 98: 12701–12705.
52. Beyer WR, Popplau D, Garten W, von Laer D, Lenz O (2003) Endoproteolytic processing of the lymphocytic choriomeningitis virus glycoprotein by the subtilase SKI-1/S1P. *J Virol* 77: 2866–2872.
53. Rojck JM, Lec AM, Nguyen N, Spiropoulou CF, Kunz S (2008) Site 1 protease is required for proteolytic processing of the glycoproteins of the South American hemorrhagic fever viruses Junin, Machupo, and Guanarito. *J Virol* 82: 6045–6051.
54. Burns JW, Buchmeier MJ (1991) Protein-protein interactions in lymphocytic choriomeningitis virus. *Virology* 183: 620–629.
55. Eichler R, Lenz O, Strecker T, Garten W (2003) Signal peptide of Lassa virus glycoprotein GP-C exhibits an unusual length. *FEBS Lett* 538: 203–206.
56. Burns JW, Buchmeier MJ (1993) Glycoproteins of the arenaviruses. In: Salvato MS, ed. *The Arenaviridae*. New York: Plenum Press. pp 17–33(35).
57. von Heijne G (1984) How signal sequences maintain cleavage specificity. *J Mol Biol* 173: 243–251.
58. York J, Romanowski V, Lu M, Nunberg JH (2004) The signal peptide of the Junin arenavirus envelope glycoprotein is myristoylated and forms an essential subunit of the mature G1–G2 complex. *J Virol* 78: 10783–10792.
59. Bowen MD, Peters CJ, Mills JN, Nichol ST (1996) Oliveros virus: a novel arenavirus from Argentina. *Virology* 217: 362–366.
60. Pinheiro FP, Shope RE, de Andrade AHP, Bensabath G, Cacios GV, et al. (1966) Amapari, a new virus of the Tacaribe group from rodents and mites of Amapa Territory, Brazil. *Proc Soc Exp Biol Med* 122: 531–535.
61. York J, Nunberg JH (2006) Role of the stable signal peptide of Junin arenavirus envelope glycoprotein in pH-dependent membrane fusion. *J Virol* 80: 7775–7780.
62. Fulhorst CE, Bowen MD, Salas RA, de Manzione NM, Duno G, et al. (1997) Isolation and characterization of pirital virus, a newly discovered South American arenavirus. *Am J Trop Med Hyg* 56: 548–553.
63. Pircher H, Moskophidis D, Rohrer U, Burki K, Hengartner H, et al. (1990) Viral escape by selection of cytotoxic T cell-resistant virus variants in vivo. *Nature* 346: 629–633.
64. Lenz O, ter Meulen J, Feldmann H, Klenk HD, Garten W (2000) Identification of a novel consensus sequence at the cleavage site of the Lassa virus glycoprotein. *J Virol* 74: 11418–11421.
65. York J, Nunberg JH (2007) A novel zinc-binding domain is essential for formation of the functional Junin virus envelope glycoprotein complex. *J Virol* 81: 13385–13391.
66. Cao W, Henry MD, Borrow P, Yamada H, Elder JH, et al. (1998) Identification of alpha-dystroglycan as a receptor for lymphocytic choriomeningitis virus and Lassa fever virus. *Science* 282: 2079–2081.
67. Spiropoulou CF, Kunz S, Rollin PE, Campbell KP, Oldstone MB (2002) New World arenavirus clade C, but not clade A and B viruses, utilizes alpha-dystroglycan as its major receptor. *J Virol* 76: 5140–5146.
68. Radoshitzky SR, Abraham J, Spiropoulou CF, Kuhn JH, Nguyen D, et al. (2007) Transferrin receptor 1 is a cellular receptor for New World hemorrhagic fever arenaviruses. *Nature* 446: 92–96.
69. Riviere Y, Ahmed R, Southern PJ, Buchmeier MJ, Oldstone MB (1985) Genetic mapping of lymphocytic choriomeningitis virus pathogenicity: virulence in guinea pigs is associated with the L RNA segment. *J Virol* 55: 704–709.
70. Lukashevich IS, Patterson J, Carrion R, Moshkoff D, Ticer A, et al. (2005) A live attenuated vaccine for Lassa fever made by reassortment of Lassa and Mopeia viruses. *J Virol* 79: 13934–13942.
71. Archer AM, Rico-Hesse R (2002) High genetic divergence and recombination in Arenaviruses from the Americas. *Virology* 304: 274–281.
72. Charrel RN, de Lamballerie X, Fulhorst CF (2001) The Whitewater Arroyo virus: natural evidence for genetic recombination among Tacaribe serocomplex viruses (family Arenaviridae). *Virology* 283: 161–166.
73. Charrel RN, Feldmann H, Fulhorst CF, Khelifa R, de Chesne R, et al. (2002) Phylogeny of New World arenaviruses based on the complete coding sequences of the small genomic segment identified an evolutionary lineage produced by intrasegmental recombination. *Biochem Biophys Res Commun* 296: 1118–1124.
74. Palacios G, Quan PL, Jabado OJ, Conlan S, Hirschberg DL, et al. (2007) Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg Infect Dis* 13: 73–81.
75. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, et al. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science* 230: 1350–1354.
76. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.
77. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, et al. (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318: 283–287.
78. Rice P, Longden I, Bleasby A (2000) EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* 16: 276–277.
79. Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5: 150–163.