Novel Virus Related to Kaposi's Sarcoma–Associated Herpesvirus from Colobus Monkey

Akshay Dhingra, Tina Ganzenmueller, Elias Hage, Nicolás M. Suárez, Kerstin Mätz-Rensing, Dimitri Widmer, Stefan Pöhlmann, Andrew J. Davison,[1] Thomas F. Schulz,[1] Artur Kaul[1]

We determined the complete genome sequence of a virus isolated from a mantled guereza that died of primary effusion lymphoma. The virus is closely related to Kaposi's sarcoma–associated herpesvirus (KSHV) but lacks some genes implicated in KSHV pathogenesis. This finding may help determine how KSHV causes primary effusion lymphoma in humans.

Kaposi's sarcoma–associated herpesvirus (KSHV) causes Kaposi sarcoma, primary effusion lymphoma, and the plasma cell variant of multicentric Castleman disease in humans (*1*). KSHV-related viruses (also known as rhadinoviruses) naturally infect New and Old World primates (*2*–*5*). Old World primate rhadinoviruses fall into 2 lineages, rhadinovirus 1 (RV1) and rhadinovirus 2 (RV2) (*2*,*6*). The RV1 lineage contains KSHV; the retroperitoneal fibromatosis–associated herpesviruses (RFHVs) identified in *Macaca nemestrina*, *M. fascicularis*, and *M. mulatta* macaques; and closely related viruses of other Old World primates (*2*,*5*). The RV2 lineage contains macaque viruses more distantly related to KSHV, such as rhesus macaque rhadinovirus, *M. nemestrina* RV2, *M. fascicularis* RV2, and Japanese macaque rhadinovirus (*2*,*7*). Complete genome sequences of the RV1 lineage viruses KSHV and RFHV of *M. nemestrina* macaques (RFHVMn), as well as of the RV2 lineage viruses rhesus macaque rhadinovirus, Japanese macaque rhadinovirus, and *M. nemestrina* RV2, have been generated from cultured viruses or directly from clinical material by conventional or high-throughput sequencing (*7*–*11*).

Apart from KSHV, all fully sequenced Old World primate rhadinoviruses have been found in primates of the genus *Macaca*, subfamily Cercopithecinae (*7*–*9*,*12*–*14*). We describe a novel rhadinovirus of the Old World primate genus *Colobus* (*14*), subfamily Colobinae, which was detected in a mantled guereza (*Colobus guereza kikuyensis*) that died of primary effusion lymphoma. The virus belongs to the RV1 lineage, together with KSHV and RFHVMn.

The Study

In 2014, a 3-year-old male mantled guereza at a zoo in Germany died suddenly after developing severe anemia (<5 g/dL hemoglobin), subcutaneous edema, and leukocytosis. A necropsy conducted at the German Primate Centre, Göttingen, Germany, led to a diagnosis of primary effusion lymphoma. Large numbers of abnormal leukocytes were found in the vascular system of several organs. The pleura pulmonalis and the pleural space were severely infiltrated with pleomorphic round cells (Figure 1, panel A) identified as CD20-positive B lymphocytes (Figure 1, panel B) with high expression of the proliferation marker Ki67 (Figure 1, panel C). Many neoplastic cells also showed typical nuclear staining with antibodies against the KSHV latent nuclear-associated antigen, suggesting infection with a related herpesvirus (Figure 1, panel D). We detected viral genomes in several organs by using PCR with a panherpesvirus primer set, a primer set specific for the virus detected in this study (colobine gammaherpesvirus 1 [CbGHV1]), or both, for the viral DNA polymerase gene. Sanger sequencing of the panherpes PCR products followed by BLAST (https://blast.ncbi.nlm.nih.gov) analysis revealed the best match to be RFHVMn. Using a commercial microarray (Simian Panel E Kit; Intuitive Biosciences, http://intuitivebio.com), we detected antibodies to lymphocryptovirus but not to simian immunodeficiency virus, simian retrovirus, herpes B virus, simian T-cell leukemia virus, measles virus, rhesus

Author affiliations: Hannover Medical School Institute of Virology, Hannover, Germany (A. Dhingra, T. Ganzenmueller, E. Hage, T.F. Schulz); German Centre for Infection Research, Hannover-Braunschweig Site, Germany (A. Dhingra, T. Ganzenmueller, E. Hage, T.F. Schulz); University Hospital Tübingen, Tübingen, Germany (T. Ganzenmueller); MRC–University of Glasgow Centre for Virus Research, Glasgow, Scotland, UK (N.M. Suárez, A.J. Davison); German Primate Center–Leibniz Institute for Primate Research, Göttingen, Germany (K. Mätz-Rensing, S. Pöhlmann, A. Kaul); Zoo Dresden GmbH, Dresden, Germany (D. Widmer); University Göttingen, Göttingen (S. Pöhlmann)

DOI: https://doi.org/10.3201/eid2508.181802

[1]These authors contributed equally to this article.

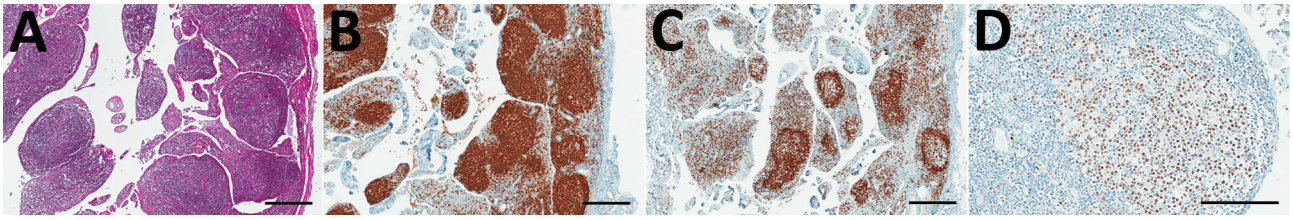


Figure 1. Identification of primary effusion lymphoma and immunohistochemical staining of primary effusion lymphoma cells with Kaposi's sarcoma–associated herpesvirus latent nuclear-associated antigen (LANA)–specific antibody. A) Diffuse infiltration of the pleura pulmonalis and pleural space with pleomorphic round cells resembling primary effusion lymphoma. Hematoxylin and eosin stain; scale bar indicates 400 μm. B) The neoplastic cells are lymphocytic cells of B cell origin. CD20 immunohistochemistry; scale bar indicates 400 μm. C) Numerous neoplastic cells express the proliferation marker Ki67. Ki67 immunohistochemistry; scale bar indicates 400 μm. D) Typical nuclear expression of a protein related to Kaposi's sarcoma–associated herpesvirus LANA in neoplastic cells. LANA immunohistochemistry; scale bar indicates 200 μm.

macaque rhadinovirus, human cytomegalovirus, or simian foamy virus (data not shown).

DNA extracted from a spleen necropsy specimen was sequenced by using an Illumina MiSeq (https://www.illumina.com). The 22,978,561 trimmed reads were depleted of host sequences by screening against the human genome sequence. The remaining 3,082,106 reads were assembled de novo into contigs, the largest of which was 126,024 bp. Assemblies of the initial trimmed reads with this sequence, followed by manual extension and incorporation of smaller contigs, resulted in a final, complete viral sequence of 133,999 bp. This essentially circular sequence consists of a unique region (U; 132,514 bp; 52% G+C) followed by a copy of a terminal repeat unit (TR; 758 bp; 84% G+C) and then by a partial copy of TR (727 bp). A total of 84,532 (0.4%) of the initial trimmed reads aligned with this sequence at an average depth of 170 reads per nucleotide. Inspection of the read alignment indicated that most genomes (85%) lack a 7,045-bp region toward the right end of U. In addition, a telomere-like tandem repeat was noted near the left end of U. A similar feature is present in the genome of RFHVMn but not KSHV.

The viral genome sequence is 51% identical to that of KSHV (137,969 bp) and 59% identical to that of RFHVMn (127,320 bp). Phylogenetic analysis of these 3 sequences with those of viruses in the RV2 lineage, using Epstein-Barr virus (a lymphocryptovirus) as the outgroup, confirmed that the novel virus clusters in the RV1 lineage with RFHVMn and KSHV (Figure 2). The novel virus was thus

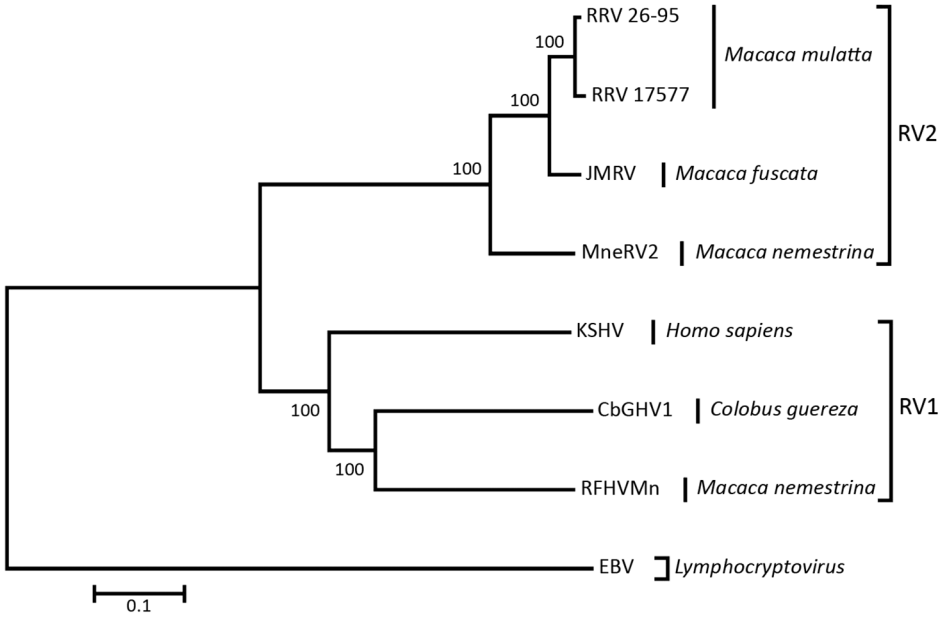


Figure 2. Nucleotide sequence–based phylogenetic analysis of the genomes of CbGHV1 and other gammaherpesviruses. The genus *Lymphocryptovirus* is represented by Epstein-Barr virus as outgroup, and the genus *Rhadinovirus* is represented by the RV1 and RV2 lineages, with host species indicated. Sequences are based on the complete U region, bootstrap values are shown as percentages, and the scale bar represents nucleotide substitutions per site. CbGHV1, colobine gammaherpesvirus 1 (KHSV-like virus isolated from a mantled guereza); EBV, Epstein-Barr virus; JMRV, Japanese macaque rhadinovirus; KSHV, Kaposi's sarcoma–associated herpesvirus; MneRV2, *Macaca nemestrina*; RV2; RFHVMn, retroperitoneal fibromatosis–associated herpesviruses of *Macaca nemestrina* macaques; RRV, rhesus macaque RV; RV, *Rhadinovirus*.

distinguished from other rhadinoviruses, and we named it colobine gammaherpesvirus 1 (CbGHV1).

We named the 78 protein-coding genes annotated in the CbGHV1 genome according to the KSHV nomenclature (Appendix Figure 1, https://wwwnc.cdc.gov/EID/article/25/8/18-1802-App1.pdf). All are located in U and have orthologs in both KSHV and RFHVMn (Appendix Table). Reanalysis of genome sequences confirmed that RFHVMn contains 82 genes, whereas KSHV contains 86 genes. Several genes first described in KSHV lack orthologs in CbGHV1 (K2, K4.2, K5, K6, K7, and K12) and RFHVMn (K5, K6, and K12). In addition, CbGHV1 lacks open reading frame (ORF) 11 (as does RFHVMn) and ORF49, and ORF2 is truncated. In comparison with RFHVMn, CbGHV1 lacks K2, K4.2, K7, and ORF49. The deletion present in most CbGHV1 genomes affects part of ORF68, all of ORF69, and part of ORF71. Values for percentage amino acid sequence identity between CbGHV1 genes and their counterparts in KSHV and RFHVMn are listed in the Appendix Table. An alignment of the KSHV, RFHVMn, and CbGHV1 latent nuclear-associated antigen (ORF73) sequences showed that all 3 contain the typical extended internal repeat region (Appendix Figure 2).

Conclusions

We identified and sequenced the complete genome of a novel KSHV-like virus (CbGHV1) from a mantled guereza. The animal died of primary effusion lymphoma, which we assume was caused by CbGHV1. CbGHV1 and its close relatives KSHV and RFHVMn cluster in the RV1 lineage. The presence of a telomere-like tandem repeat near the left end of U in the CbGHV1 and RFHVMn genomes suggests that an ancestral virus may have been integrated into the host genome, and its persistence suggests that these viruses may retain the ability to integrate.

The CbGHV1 genome contains all genes that are conserved in all members of the family *Herpesviridae*. Orthologs of 8 KSHV genes (ORF11, K2, K4.2, K5, K6, K7, ORF49, and K12) are absent from CbGHV1. In KSHV, some of these genes, such as K2 and K12, encode proteins (vIL6 and kaposin, respectively) that have been linked to viral pathogenesis. Their absence from CbGHV1 suggests that they may not be needed for the development of primary effusion lymphoma. Because ORF10 and ORF11 are related and may have arisen by duplication from an ancestral deoxyuridine triphosphatase gene (*15*), their functions may overlap. ORF49 is a cofactor of the KSHV lytic cycle activator Rta (ORF50) and may not be required for the function of CbGHV1 Rta. It is likely that the viral subpopulation lacking all or part of ORF68, ORF69, and ORF71, which encode essential proteins involved in packaging of viral DNA into capsids, egress of capsids from the nucleus, and inhibition of apoptosis, represents a replication-defective deletion mutant generated in the animal investigated. Of note, CbGHV1 has also been identified by PCR in an older female sibling guereza with Kaposi sarcoma–like disease (*16*). The latest offspring of this female guereza were also positive for CbGHV1 but did not show clinical signs. Discovery of CbGHV1 in multiple animals and the determination of its genome sequence may inform future studies of the pathogenesis of primary effusion lymphoma and Kaposi sarcoma, including how KSHV causes primary effusion lymphoma in humans.

This work was supported by the Deutsche Forschungsgemeinschaft Collaborative Research Centre 900 (grant no. 158989968–SFB900 core project Z1) and the UK Medical Research Council (grant no. MC_UU_12014/3). Akshay Dhingra was supported by the Infection Biology graduate program of Hannover Biomedical Research School.

About the Author

Mr. Dhingra is a PhD candidate at the Institute of Virology, Hannover Medical School. His research interest is the diversity and evolution of DNA viruses in vivo.

References

1. Cesarman E. Gammaherpesvirus and lymphoproliferative disorders in immunocompromised patients. Cancer Lett. 2011; 305:163–74. http://dx.doi.org/10.1016/j.canlet.2011.03.003
2. Bruce AG, Bielefeldt-Ohmann H, Barcy S, Bakke AM, Lewis P, Tsai CC, et al. Macaque homologs of EBV and KSHV show uniquely different associations with simian AIDS-related lymphomas. PLoS Pathog. 2012;8:e1002962. http://dx.doi.org/10.1371/journal.ppat.1002962
3. Albrecht JC, Nicholas J, Biller D, Cameron KR, Biesinger B, Newman C, et al. Primary structure of the herpesvirus saimiri genome. J Virol. 1992;66:5047–58.
4. Rose TM, Strand KB, Schultz ER, Schaefer G, Rankin GW Jr, Thouless ME, et al. Identification of two homologs of the Kaposi's sarcoma-associated herpesvirus (human herpesvirus 8) in retroperitoneal fibromatosis of different macaque species. J Virol. 1997;71:4138–44.
5. Greensill J, Sheldon JA, Renwick NM, Beer BE, Norley S, Goudsmit J, et al. Two distinct gamma-2 herpesviruses in African green monkeys: a second gamma-2 herpesvirus lineage among Old World primates? J Virol. 2000;74:1572–7. http://dx.doi.org/10.1128/JVI.74.3.1572-1577.2000
6. Schultz ER, Rankin GW Jr, Blanc MP, Raden BW, Tsai CC, Rose TM. Characterization of two divergent lineages of macaque rhadinoviruses related to Kaposi's sarcoma-associated herpesvirus. J Virol. 2000;74:4919–28. http://dx.doi.org/10.1128/JVI.74.10.4919-4928.2000
7. Estep RD, Hansen SG, Rogers KS, Axthelm MK, Wong SW. Genomic characterization of Japanese macaque rhadinovirus, a novel herpesvirus isolated from a nonhuman primate with a spontaneous inflammatory demyelinating disease. J Virol. 2013;87:512–23. http://dx.doi.org/10.1128/JVI.02194-12
8. Bruce AG, Ryan JT, Thomas MJ, Peng X, Grundhoff A, Tsai CC, et al. Next-generation sequence analysis of the genome of RFHVMn, the macaque homolog of Kaposi's sarcoma

(KS)-associated herpesvirus, from a KS-like tumor of a pig-tailed macaque. J Virol. 2013;87:13676–93. http://dx.doi.org/10.1128/JVI.02331-13
9. Bruce AG, Thouless ME, Haines AS, Pallen MJ, Grundhoff A, Rose TM. Complete genome sequence of pig-tailed macaque rhadinovirus 2 and its evolutionary relationship with rhesus macaque rhadinovirus and human herpesvirus 8/Kaposi's sarcoma-associated herpesvirus. J Virol. 2015;89:3888–909. http://dx.doi.org/10.1128/JVI.03597-14
10. Olp LN, Jeanniard A, Marimo C, West JT, Wood C. Whole-genome sequencing of Kaposi's sarcoma-associated herpesvirus from Zambian Kaposi's sarcoma biopsy specimens reveals unique viral diversity. J Virol. 2015;89:12299–308. http://dx.doi.org/10.1128/JVI.01712-15
11. Russo JJ, Bohenzky RA, Chien MC, Chen J, Yan M, Maddalena D, et al. Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). Proc Natl Acad Sci U S A. 1996;93:14862–7. http://dx.doi.org/10.1073/pnas.93.25.14862
12. Searles RP, Bergquam EP, Axthelm MK, Wong SW. Sequence and genomic analysis of a Rhesus macaque rhadinovirus with similarity to Kaposi's sarcoma-associated herpesvirus/human herpesvirus 8. J Virol. 1999;73:3040–53.
13. Alexander L, Denekamp L, Knapp A, Auerbach MR, Damania B, Desrosiers RC. The primary sequence of rhesus monkey rhadinovirus isolate 26-95: sequence similarities to Kaposi's sarcoma-associated herpesvirus and rhesus monkey rhadinovirus isolate 17577. J Virol. 2000;74:3388–98. http://dx.doi.org/10.1128/JVI.74.7.3388-3398.2000
14. Groves CP, Wilson DE, Reeder DM. Mammal species of the world: a taxonomic and geographic reference. 3rd ed. Baltimore (MD): Johns Hopkins University Press; 2005. p. 161–72.
15. Davison AJ, Stow ND. New genes from old: redeployment of dUTPase by herpesviruses. J Virol. 2005;79:12880–92. http://dx.doi.org/10.1128/JVI.79.20.12880-12892.2005
16. Grewer A, Bleyer M, Mätz-Rensing K, Hahn A, Rüggeberg T, Babaryka G, et al. Kaposi sarcoma in mantled guereza. Emerg Infect Dis. 2019;25:1552–5.

Address for correspondence: Thomas F. Schulz, Hannover Medical School, Institute of Virology, Carl-Neuberg-Str. 1, 30625 Hannover, Lower Saxony, Germany; email: schulz.thomas@mh-hannover.de; Andrew J. Davison, MRC–University of Glasgow Centre for Virus Research, Sir Michael Stoker Bldg, 464 Bearsden Rd, Glasgow, Scotland G61 1QH, UK; email: andrew.davison@glasgow.ac.uk

Article DOI: https://doi.org/10.3201/eid2508.181802

Novel Virus Related to Kaposi's Sarcoma–Associated Herpesvirus from Colobus Monkey

Appendix

Materials and Methods

Animal and clinical samples

A three-year-old male mantled guereza (*Colobus guereza kikuyensis*) that had been born and housed in Dresden Zoo died suddenly. A necropsy, including histopathologic, immunohistologic and other investigations of several organs, was carried out at the German Primate Center.

Histopathologic examination

Necropsy specimens were prepared as formalin-fixed, paraffin-embedded sections and analysed by hematoxylin and eosin staining. Immunohistochemistry was performed on sections using primary antibodies against human Ki67 (mouse monoclonal antibody, clone MIB-1, DakoCytomation, Hamburg, Germany, 1:50), CD20 (mouse monoclonal anti-human, clone L26, DakoCytomation, 1:300), CD3 (rabbit polyclonal antibody, DakoCytomation, 1:50) and KSHV LANA (rat monoclonal antibody, clone LN35, Abcam, UK, 1:10 recognizes the LANA EEPEPE epitope), respectively, and the streptavidin-biotin-complex method (DAB Map kit, Roche Diagnostics, Germany) in an automated immunostaining system (Discovery XT, Roche).

DNA extraction and partial sequencing

Total DNA was isolated from tissue samples (First-DNA All-Tissue kit, GEN-IAL, Germany). PCR was performed on DNA from lung and spleen samples using the published pan-herpes PCR primer sets DFA, ILK and KG1 (*1*). PCR products extracted from agarose gels were analysed by Sanger sequencing.

Quantitative CbGHV1 specific PCR

Quantitative PCR was conducted in triplicate and repeated three times using 10 ng extracted DNA in 3 mM $MgCl_2$, 0.4 mM deoxynucleoside triphosphates, 0.266 μM probe, 0.6 μM (each) sense and antisense primers, 1 x PCR buffer (Qiagen, Germany) and 0.25 μl HotStarTaq DNA polymerase (Qiagen). The thermal profile was 95 °C for 15 min followed by 45 cycles consisting of 15 s at 95 °C and 1 min at 60 °C on a Rotor-Gene Q (Qiagen). Primers and probe were based on the CbGHV1 DNA polymerase catalytic subunit (ORF9) gene sequence (5´-CCGAGACAGTAACCCTCCAA-3´, 5´-TTAGCAGGCAGGCTAAGTGT-3´, and 5´FAM-TGGCTTCCACGAAGACCTGTGACT-3´BHQ-1).

Genome sequencing

Sequencing libraries were prepared from DNA extracted from a spleen sample by using a KAPA library preparation kit (KAPA Biosystems, USA). Fragments were generated by sonication, end-repaired, A-tailed, ligated to the NEBnext Illumina adaptor (New England BioLabs, USA) and amplified by PCR using a KAPA HiFi real-time library amplification kit on an ABI 7500 real-time cycler (Applied Biosystems, USA). After quality control using a Qubit 2.0 fluorometer (Invitrogen, USA) and a Bioanalyzer (Agilent Technologies, USA), sequencing was performed on an Illumina MiSeq using a v. 3 reagent kit (Illumina), generating a dataset of 300 nucleotide paired-end reads.

Host sequences were removed by mapping the reads to the UCSC hg19 human reference genome. The remaining unmapped reads were quality-filtered (FastQC v. 0.11.5), and adapter sequences were removed (Trim Galore v. 0.4 (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore). Consensus sequences were derived *de novo* (SPAdes v. 3.10.1 (*2*)), and read assemblies were generated and checked manually (CLC genomics workbench v. 9 (Qiagen)) or generated using Bowtie 2 v. 2.3.1 (*3*) and visualized (Tablet v. 1.17.08.17 (*4*)). The viral genome sequence was annotated by comparison with the KSHV and RFHVMn genome sequences (Geneious v. 11.1.3 (*5*). Using standard bioinformatics tools, searches were also conducted for potentially novel genes that had been missed in previous analyses of the KSHV and RFHVMn sequences.

Sequence alignments and phylogenetic analyses

Nucleotide sequence alignments of the genome sequences of CbGHV1, KSHV (AF148805), RFHVMn (KF703446), EBV (NC_007605), RRV strain 26-95 (AF210726), RRV strain 17577 (AY528864), JMRV (AY528864) and MneRV2 (KP265674) were constructed using MAFFT v. 7 (*6*). Phylogenetic analyses were carried out using MEGA v. 7 (*7*), employing the neighbor-joining method with 1000 bootstrap replicates. Amino acid sequence alignments were constructed for individual genes using Bioedit v. 7.2.0 or Geneious v. 11.1.3 (*5*) to calculate percentage identity.

Genome sequence accession number

The CbGHV1 genome sequence was deposited in NCBI GenBank (accession number MH932584).

References

1. Chmielewicz B, Goltz M, Lahrmann KH, Ehlers B. Approaching virus safety in xenotransplantation: a search for unrecognized herpesviruses in pigs. Xenotransplantation. 2003;10:349–56. PubMed http://dx.doi.org/10.1034/j.1399-3089.2003.02074.x

2. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012:19:455-77.

3. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357–9. PubMed http://dx.doi.org/10.1038/nmeth.1923

4. Milne I, Stephen G, Bayer M, Cock PJ, Pritchard L, Cardle L, et al. Using Tablet for visual exploration of second-generation sequencing data. Brief Bioinform. 2013;14:193–202. PubMed http://dx.doi.org/10.1093/bib/bbs012

5. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28:1647–9. PubMed http://dx.doi.org/10.1093/bioinformatics/bts199

6. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–66. PubMed http://dx.doi.org/10.1093/nar/gkf436

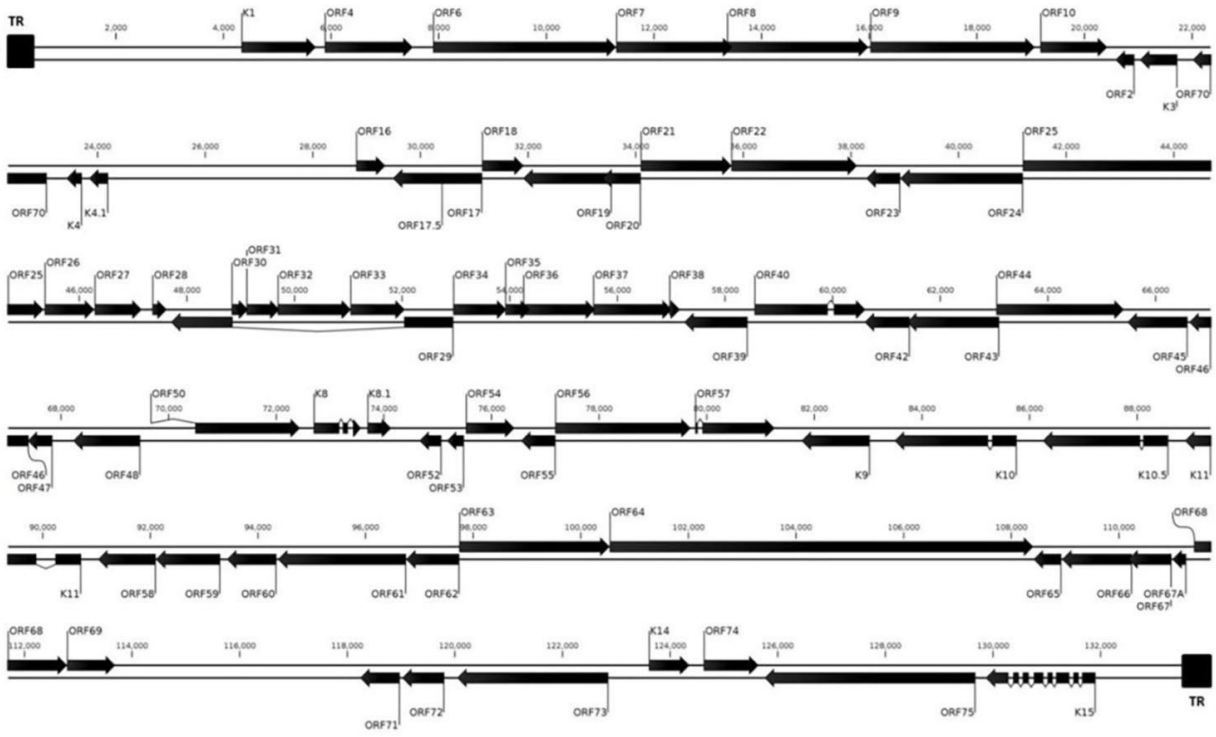
7. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;33:1870–4. PubMed http://dx.doi.org/10.1093/molbev/msw054

Appendix Table. Functions of CbGHV1 genes and amino acid sequence comparisons with orthologous genes in KSHV, RFHVMn, and RRV*

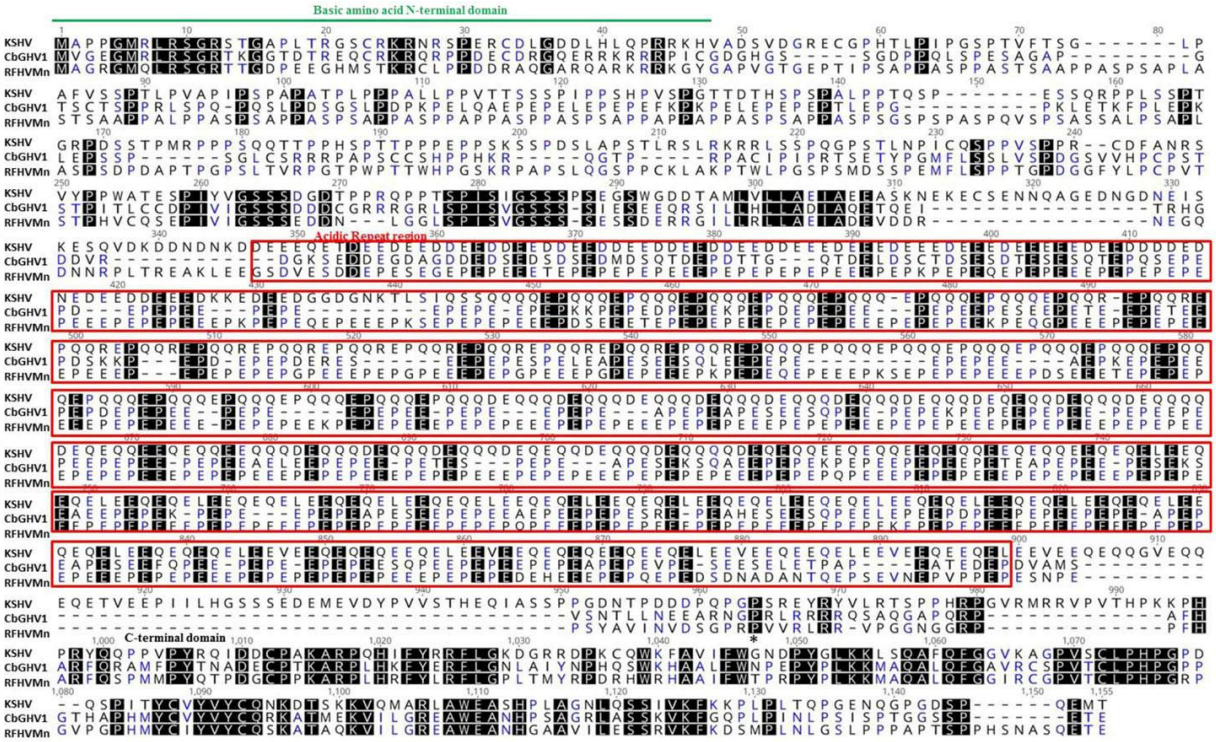
Gene	Protein product	Identity to RFHVMn (%)	Identity to KSHV (%)	Identity to RRV (%)
K1	Membrane glycoprotein K1	36.3	13.2	18.8
ORF4	Complement control protein	44.4	37.7	25.3
ORF6	Single-stranded DNA-binding protein	78.6	72.9	61.2
ORF7	DNA packaging terminase subunit 2	67.4	62.9	51.4
ORF8	Envelope glycoprotein gB	71.7	66.7	60.1
ORF9	DNA polymerase catalytic subunit	81.0	73.9	66.7
ORF10	Protein G10	56.7	40.8	32.6
ORF2	Dihydrofolate reductase	41.8	37.5	38.3
K3	E3 ubiquitin ligase MIR1	44.4	31.1	19.4
ORF70	Thymidylate synthase	69.8	60.7	62.6
K4	CC chemokine vCCLl2	57.4	54.3	17.0
K4.1	CC chemokine vCCL3	58.3	40.3	19.0
ORF16	Apoptosis regulator G16	53.9	46.3	37.0
ORF17	Capsid maturational protease	60.3	49.8	41.5
ORF17.5	Capsid scaffold protein	49.4	33.8	25.6
ORF18	Protein UL79	71.9	65.4	56.8
ORF19	DNA packaging tegument protein UL25	71.5	60.7	49.6
ORF20	Nuclear protein UL24	61.0	49.6	37.2
ORF21	Thymidine kinase	60.6	49.0	36.0
ORF22	Envelope glycoprotein gH	61.1	43.6	36.0
ORF23	Protein UL88	43.9	38.4	24.6
ORF24	Virion protein UL87	70.2	66.5	53.7
ORF25	Major capsid protein	88.9	82.4	73.5
ORF26	Capsid triplex subunit 2	82.3	77.4	59.0
ORF27	Envelope glycoprotein 48	51.4	42.2	25.3
ORF28	Envelope glycoprotein 150	47.3	40.8	21.5
ORF29	DNA packaging terminase subunit 1	78.1	73.7	57.3
ORF30	Protein UL91	62.3	58.8	33.3
ORF31	Protein UL92	75.9	70.1	47.3
ORF32	DNA packaging tegument protein UL17	55.6	44.1	37.4
ORF33	Tegument protein UL16	75.5	61.2	42.9
ORF34	Protein UL95	69.6	61.8	47.4
ORF35	Tegument protein UL14	65.3	53.3	30.9
ORF36	Tegument serine/threonine protein kinase	72.1	61.7	43.1
ORF37	Deoxyribonuclease	82.9	70.6	65.7
ORF38	Myristylated tegument protein	59.4	53.1	43.9
ORF39	Envelope glycoprotein gM	75.6	63.0	58.9
ORF40	Helicase-primase subunit	51.6	43.3	30.0
ORF42	Tegument protein UL7	70.8	60.6	46.0
ORF43	Capsid portal protein	80.1	74.7	60.3
ORF44	Helicase-primase Helicase subunit	83.5	76.1	67.0
ORF45	Tegument protein G45	52.1	35.1	22.6
ORF46	Uracil-DNA glycosylase	77.6	67.1	54.2
ORF47	Envelope glycoprotein gL	57.3	39.6	31.9
ORF48	Tegument protein G48	49.0	32.3	29.2
ORF50	Protein Rta	54.1	49.7	37.5
K8	Protein Zta	54.4	29.7	12.9
K8.1	Glycoprotein gp350	28.9	21.0	12.4
ORF52	Virion protein G52	66.9	53.0	42.6
ORF53	Envelope glycoprotein gN	56.7	47.3	45.9
ORF54	Deoxyuridine triphosphatase	67.6	48.1	37.8
ORF55	Tegument protein UL51	69.4	64.8	51.9
ORF56	Helicase-primase primase subunit	66.9	58.6	50.2
ORF57	Mutifunctional expression regulator	64.5	55.8	40.4
K9	Interferon regulatory factor 1	53.6	36.9	17.7
K10	Interferon regulatory factor 4	38.6	24.8	NA
K10.5	Interferon regulatory factor 3	34.1	24.4	NA
K11	Interferon regulatory factor 2	35.8	26.1	NA
ORF58	Envelope protein UL43	62.9	49.6	36.7

Gene	Protein product	Identity to RFHVMn (%)	Identity to KSHV (%)	Identity to RRV (%)
ORF59	DNA polymerase processivity subunit	65.8	57.8	47.1
ORF60	Ribonucleotide reductase subunit 2	84.9	75.7	70.5
ORF61	Ribonucleotide reductase subunit 1	80.1	66.9	61.3
ORF62	Capsid triplex subunit 1	76.2	64.4	51.8
ORF63	Tegument protein UL37	57.3	47.3	37.4
ORF64	Large tegument protein	56.6	46.8	36.2
ORF65	Small capsid protein	55.1	45.2	35.7
ORF66	Protein UL49	62.6	57.3	42.7
ORF67	Nuclear egress membrane protein	75.8	59.4	56.6
ORF67A	DNA packaging protein UL33	72.6	60.7	57.0
ORF68	DNA packaging protein UL32	60.0	59.1	49.2
ORF69	Nuclear egress lamina protein	78.7	66.2	60.4
ORF71	Apoptosis regulator FLIP	50.8	45.5	35.6
ORF72	Cyclin	46.3	45.6	37.5
ORF73	Nuclear antigen LANA-1	46.3	26.3	14.6
K14	Glycoprotein CD200	57.1	37.5	31.2
ORF74	Membrane protein G74	65.4	56.2	42.4
ORF75	Tegument protein G75	75.6	60.5	42.1
K15	Membrane protein K15	47.1	20.5	11.6

*CbGHV1, KHSV-like virus isolated from a mantled guereza; KSHV, Kaposi sarcoma–associated herpesvirus; LANA, latent nuclear-associated antigen; NA, not applicable; ORF, open reading frame; RFHVMn, retroperitoneal fibromatosis–associated herpesviruses identified in *Macaca nemestrina* macaques; RRV, rhesus macaque rhadinovirus



Appendix Figure 1. CbGHV1 genome map. The genome is depicted as U flanked at each end by a single copy of TR, although variable numbers of TR are likely present at each end. Protein-coding regions are represented as black arrows. The scale is in base pairs (bp).



Appendix Figure 2. Amino acid sequence alignment of LANA (ORF73) for members of the RV1 lineage. Identical residues in three sequences are highlighted, and identical residues in two sequences are in blue font. The N-terminal domain is indicated in green font and the acidic repeat region in red font, and the C-terminal domain starts with an asterisk at residue 964.