Medical Microbiology and Immunology (2019) 208:109–129
https://doi.org/10.1007/s00430-018-0565-y

ORIGINAL INVESTIGATION

Macaca arctoides gammaherpesvirus 1 (strain herpesvirus Macaca arctoides): virus sequence, phylogeny and characterisation of virus-transformed macaque and rabbit cell lines

Andi Krumbholz[1] · Janine Roempke[1] · Thomas Liehr[2] · Marco Groth[3] · Astrid Meerbach[4] · Michael Schacke[4] · Gregor Maschkowitz[1] · Helmut Fickenscher[1] · Wolfram Klapper[5] · Andreas Sauerbrei[4] · Peter Wutzler[4] · Roland Zell[4]

Received: 20 July 2018 / Accepted: 27 September 2018 / Published online: 5 October 2018

Abstract

Herpesvirus Macaca arctoides (HVMA) has the propensity to transform macaque lymphocytes to lymphoblastoid cells (MAL-1). Inoculation of rabbits with cell-free virus-containing supernatant resulted in the development of malignant lymphomas and allowed isolation of immortalised HVMA-transformed rabbit lymphocytes (HTRL). In this study, the HVMA genome sequence (approx. 167 kbp), its organisation, and novel aspects of virus latency are presented. Ninety-one open reading frames were identified, of which 86 were non-repetitive. HVMA was identified as a *Lymphocryptovirus* closely related to Epstein–Barr virus, suggesting the designation as 'Macaca arctoides gammaherpesvirus 1' (MarcGHV-1). In situ lysis gel and Southern blot hybridisation experiments revealed that the MAL-1 cell line contains episomal and linear DNA, whereas episomal DNA is predominantly present in HTRL. Integration of viral DNA into macaque and rabbit host cell genomes was demonstrated by fluorescence in situ hybridisation on chromosomal preparations. Analysis of next-generation sequencing data confirmed this finding. Approximately 400 read pairs represent the overlap between macaque and MarcGHV-1 DNA. Both, MAL-1 cells and HTRL show characteristics of a polyclonal tumour with B- and T-lymphocyte markers. Based on analysis of viral gene expression and immunohistochemistry, the persistence of MarcGHV-1 in MAL-1 cells resemble the latency type III, whereas the expression pattern observed in HTRL was more comparable with latency type II. There was no evidence of the presence of STLV-1 proviral DNA in MAL-1 and HTRL. Due to the similarity to EBV-mediated cell transformation, MarcGHV-1 expands the available in vitro models by simian and rabbit cell lines.

Keywords Lymphocryptovirus · Old World monkey · Genome sequence · Lymphocytes · Tumour

Introduction

At present, the subfamily *Gammaherpesvirinae* is divided into four genera named *Macavirus, Percavirus, Rhadinovirus*, and *Lymphocryptovirus*. The *Human gammaherpesvirus 4*, better known as Epstein–Barr virus (EBV),

Andi Krumbholz and Janine Roempke contributed equally to this study.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00430-018-0565-y) contains supplementary material, which is available to authorized users.

✉ Andi Krumbholz
krumbholz@infmed.uni-kiel.de

[1] Institute of Infection Medicine, University of Kiel and University Medical Center Schleswig-Holstein, Brunswiker Strasse 4, 24105 Kiel, Germany

[2] Institute of Human Genetics, Jena University Hospital, Friedrich Schiller University, Jena, Germany

[3] CF DNA sequencing, Leibniz Institute on Aging-Fritz Lipmann Institute (FLI), Jena, Germany

[4] Division of Experimental Virology, Institute of Medical Microbiology, Jena University Hospital, Friedrich Schiller University, Jena, Germany

[5] Department of Pathology, Hematopathology Section and Lymph Node Registry, University of Kiel and University Medical Center Schleswig-Holstein, Kiel, Germany

represents the type species of the genus *Lymphocryptovirus* which also includes gammaherpesviruses of different primates, such as *Callitrichine gammaherpesvirus 3, Cercopithecine gammaherpesvirus 14, Gorilline gammaherpesvirus 1, Macacine gammaherpesvirus 4, Panine gammaherpesvirus 1, Papiine gammaherpesvirus 1* and *Pongine gammaherpesvirus 2* (ICTV Master Species List 2017 v1.0: https://talk.ictvonline.org/files/master-species-lists/m/msl/7185).

Similar to other gammaherpesviruses, EBV is lymphotropic but also capable of undergoing lytic replication in epithelial cells [1]. The primary infection of B-lymphocytes by EBV in early childhood is mostly asymptomatic, whereas the clinical picture of mononucleosis is frequently observed in adolescents or adults [2, 3]. After primary infection, EBV establishes a life-long latency in subsets of memory B-lymphocytes [4]. Maintenance of latency is achieved by expression of up to nine viral genes encoding latency proteins, but also by viral RNA, altogether expediting the development of the naïve B-lymphocyte to memory B-cell [5]. Epstein–Barr virus infection is etiologically linked with the occurrence of lymphomas (Hodgkin's and non-Hodgkin's lymphoma, endemic Burkitt lymphoma, etc.), lymphoma-like diseases (post-transplant lymphoproliferative disease) and solid epithelial tumours (nasopharyngeal and gastric carcinoma) [1, 5–7]. Furthermore, in solid organ transplant recipients, EBV reactivation has been associated with graft rejection [8].

Since herpesvirus phylogeny resembles host phylogeny, it was assumed that all primate lymphocryptoviruses (LCV) have co-evolved with their hosts [9–11]. The *Macacine gammaherpesvirus 4* (McHV-4; also known as cercopithecine herpesvirus 15 or rhesus-EBV-like virus [12]), isolated from lymphoma cells of the Old World primate *Macaca mulatta*, can serve as an excellent example for this hypothesis since McHV-4 exhibits an EBV-like genome organisation as well as an EBV-like cell tropism and pathogenesis [10, 13–15]. The similarity of lytic proteins of EBV and McHV-4 ranges between 70–90% amino acid (aa) identity, while similarity of latent proteins is given with 30–80% [15]. Particularly those genes which are associated with cell transformation are more divergent but still have similar functions [16]. Recently, the complete genome sequence of a LCV isolated from lymphoblastic cells of *Macaca fascicularis* (cynomolgus macaque LCV) was published [9]. This LCV shares the same genomic organisation of EBV and McHV-4 with the highest degree of similarity to McHV-4. Again, a lower similarity was observed for latency genes [9]. In contrast, the New World primate *Callitrichine gammaherpesvirus 3* (CalHV-3), isolated from lymphoma of *Callithrix jacchus*, differs remarkably in its genetic composition from EBV and Old World LCV while it has similar biological properties [17]. Complete sequence data are available for EBV and the three above-mentioned primate LCVs; further gene sequences have been published from additional non-human primate LCVs [9, 10, 17].

Previously, the biological properties of herpesvirus Macaca arctoides (HVMA) isolated from MAL-1 cell line have been studied: this virus was derived from peripheral blood lymphocytes of the stump-tailed macaque, *Macaca arctoides* [18] and is unable to infect human cord blood cells but can induce lymphomas in rabbits [19–23]. Infected rabbits produce antibodies which are cross-reactive against early and late proteins of EBV. Thus, HVMA was suggested to be a LCV with EBV-like pathogenesis [22]. This virus was also used in our laboratory to transform rabbit lymphocytes (HTRL, Michael Schacke, unpublished); however, sequence information was not available, which is an absolute prerequisite for deeper understanding of the biological properties and taxonomic assignment of HVMA. In addition, it is controversial whether HVMA or the simian T-lymphotropic virus 1 (STLV-1) contributed to the cellular transformation [23]. Both MAL-1 cells and the transformed rabbit cells were found by PCR to contain provirus DNA of STLV-1 [24]. Serological studies, however, did not confirm the presence of STLV-1 proteins [22].

In this report, the HVMA genome sequence and its organisation was determined and compared to EBV. Moreover, the HVMA infection status and the cell-type-specific markers of MAL-1 and HTRL were examined using in situ lysis gel, analysis of RNA transcripts, chromosome analysis, in situ hybridisation (ISH as well as fluorescence ISH, FISH), and immunohistochemistry (IHC). Finally, we reassessed the presence of STLV-1 proviral DNA in both cell lines. The results suggest that HVMA-transformed cell lines may be used as expedient, additional model systems for certain aspects of transformation by EBV and EBV-like viruses.

Materials and methods

Cell lines and cell culture

The MAL-1 cell line originates from a malignant lymphoma of *Macaca arctoides* and is productively infected with HVMA [18, 22]. These cells were obtained from the former Institute of Experimental Pathology and Therapy of the USSR Academy of Medical Sciences, Sukhumi, Georgia. The second HVMA genome-carrying cell line was derived from virus-transformed rabbit lymphocytes (HTRL) at the former Institute of Virology and Antiviral Therapy in Jena (Michael Schacke, unpublished). The Raji cell line (ATCC® CCL-86™) represents a human lymphoblast-like cell line from a Burkitt lymphoma (BL) [25, 26] and served as an EBV genome-carrying positive control. Ramos cells (ATCC® CRL-1596™), which are EBV-free BL cells [27], were used as a negative control. All cell lines were maintained in RPMI 1640 supplemented with 10% foetal

calf serum, 1% L-glutamine and 1% penicillin–streptomycin (all GE Healthcare Bio-Sciences Austria GmbH, Pasching, Austria). Phorbolester (TPA, Sigma-Aldrich Chemie GmbH, Munich, Germany), 100 ng per ml cell culture medium, was used for induction of lytic viral replication cycle as described previously [28].

DNA extraction, sequencing and analysis of Illumina sequencing data

DNA was extracted with the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. For library preparation, 1 µg DNA of MAL-1 cells was introduced into TruSeq DNA LT Sample Preparation Kit (Illumina Inc., San Diego, CA, USA) according to the manufacturer's protocol. Sequencing was conducted on an Illumina HiSeq2000 platform (2 × 100 cycle paired-end mode, one lane). Prior to assembly, herpesvirus-specific reads were identified by mapping the resulting read pairs under low stringency conditions to a collection of 50 full-length and partial LCV genomes. Subsequent de novo assembly was done using CLC assembly cell (Qiagen/CLC Bio) using all read pairs of which at least one mate was mappable in the approach mentioned. Contigs were aligned manually to the published genome sequence of McHV-4 (Acc. No. NC_006146.1). Large gaps were closed by Illumina sequencing and de novo assembly of cloned amplicons: for this, PCR fragments were generated using HVMA-specific oligonucleotides and the Long PCR Enzyme Mix kit, Phusion Hot Start II High-Fidelity DNA-Polymerase kit or the DreamTaq PCR Master Mix kit (all ThermoFisher Scientific, Dreieich, Germany) according to manufacturer's instructions. PCR products were separated by agarose gel electrophoresis and gel purified using the NucleoSpin Gel and PCR Clean up (MACHEREY–NAGEL GmbH & Co. KG, Düren, Germany) or GeneJET PCR Purification Kit (ThermoFisher Scientific). Purified amplicons were cloned into pJET1.2 plasmid or pCR2.1-TOPO applying the CloneJET PCR or TOPO® TA Cloning® Kits (ThermoFisher Scientific) and competent *E.coli* cells. A mixture of plasmids (total 1 µg) was sequenced using the MiSeq platform (2 × 250 cycle paired-end module), library preparation: TruSeq DNA LT Sample Preparation Kit. Read pairs containing only plasmid-derived sequences were excluded before performing de novo assembly using CLC assembly cell. Small sequence gaps were closed by conventional Sanger sequencing with the BigDye® Terminator v3.1 Cycle Sequencing Kit and a 3730xl DNA Analyzer (both ThermoFisher Scientific). For identification of HVMA ORFs, the viral genome organizer 3.1 (University of Victoria) was applied by comparison of the HVMA aa sequences to that of known EBV proteins.

Read pairs that comprised both macacine and HVMA sequences were considered as an indication of virus integration into the host genome. To identify such reads/events, Illumina read pairs were mapped to the macaque genome (rheMac8) using bowtie2 (v2.2.9) [29]. Read pairs of which only one mate was mappable were extracted. The non-mappable mates were subsequently mapped to the HVMA genome. The reciprocal approach (mapping to the HVMA genome followed by mapping to the macaque genome macFas5) was conducted to confirm the results.

Search for the proviral *tax* gene

Nucleic acid preparations from MAL-1 and HTRL were repeatedly tested in a nested PCR procedure for the presence of the proviral *tax* gene of STLV-1 [30] using the DreamTaq PCR Master Mix kit. A sample containing proviral HTLV-1 DNA was used as positive control. By quantitative PCR [31, 32], this sample was found to carry approx. 300 copies per µl. PCR products were separated on a 2% agarose gel. Specificity of amplification was demonstrated by Sanger sequencing. Illumina sequence data were also checked for reads mapping to a set of 377 (sub-) genomic STLV-1 and STLV-3 sequences using bowtie2 (v2.2.9).

Transcription analysis

RNA was extracted with the RNeasy Kit (Qiagen) according to the manufacturer's instructions and quantified spectrophotometrically. Then, libraries were prepared by introducing 1 µg of total RNA into Illumina's TruSeq stranded mRNA library preparation kit following the manufacturer's description. Sequencing was done using the Illumina HiSeq2500 (one lane; 50 bp single-end, high-output mode). The resulting reads were mapped to available mRNA sequences of genes characteristic for B- (CD19, CD20, CD21 and CD79α) and T-lymphocytes (CD2, CD3ε/γ, CD4, CD8α/β) but also for granulocytes, monocytes and macrophages (CD14, CD93, CD163) of macaques and rabbits using bowtie2 (v2.2.9). Furthermore, reads were mapped to latency and lytic markers of HVMA (EBER1, -2, EBNA-1, -2, -3A, -3B, -3C, -LP, LMP-1, LMP-2A, -2B, BZLF1/ZEBRA and BSLF2/BMLF1).

Phylogenetic analysis

The aa sequences of the polymerase of MarcGHV-1 and 99 other gammaherpesviruses were aligned together with five reference sequences of the subfamilies *Alpha-* and *Betaherpesvirinae* applying the MUSCLE algorithm [33, 34] implemented in MEGA 6.06 [35]. The substitution model was determined with the "Find Best DNA/Protein Models" option and used for tree reconstruction. The DNA

sequences of 42 concatenated core genes [36] obtained from MarcGHV-1, 21 further gammaherpesviruses, and four alpha- and betaherpesviruses were aligned applying the MUSCLE algorithm. Then, the best nucleotide substitution model was determined as described above. Tree inference was conducted with MEGA 6.06 [35] using the maximum likelihood (ML) method. Statistical support of tree topology was obtained by bootstrap analysis (1000 replications).

In situ lysis gel and Southern blot

The electrophoretic DNA separation technique initially described by Gardella et al. was used to distinguish linear and episomal DNA of HVMA [37]. The in situ lysis gel was prepared as described previously [37], but proteinase K (Qiagen) was used for lysis instead of pronase. Each gel slot was loaded with 1×10^6 cells suspended in 40 µl sample buffer (20% Ficoll; 0.01% bromphenol blue and 50 µg/ml RNaseA in buffer TBE). Electrophoresis was performed in two steps at 4 °C: first, cell lysis proceeded at 0.8 V/cm (15 V) for 3 h, followed by separation of DNA at 4.5 V/cm (83 V) for 20 h. After separation, DNA was blotted onto a positively charged nylon membrane (Roche, Mannheim, Germany) and hybridised with a probe specific for HVMA DNA. The 1537 bp HVMA probe is part of the BSLF1- and BSRF1-ORF, shows 86.9% homology to EBV (HHV-4 B95-8, GenBank: AJ507799, NC_007605, nt 74,025–75,543), and was amplified with the Long PCR Enzyme Mix kit and HVMA-specific primers (data not shown).

This DNA probe was labelled with digoxigenin applying the DIG-High Prime DNA Labeling and Detection Starter Kit II (Roche). Hybridisation was performed over night at 42.5 °C with DIG Easy Hyb granules (Roche) and chemiluminescence was detected with LAS-3000 Image Reader (Fujifilm Europe, Düsseldorf, Germany).

Detection of EBER and EBV latency proteins

Immunohistochemistry and EBER detection was done on 1–2 µm sections of paraffin-embedded MAL-1 and HTRL. First, cultured cells were pelleted by centrifugation (800 rpm; 10 min) and washed twice with PBS. Then, the cells were re-suspended for 20 min in 4.5% formalin followed by two centrifugation and washing steps with PBS. Thereafter, the pellet was re-suspended in 2 ml isopropanol together with three droplets of protein glycerol followed by centrifugation. The remaining pellet was brought on filtre paper and used for paraffin-embedding applying standard histological procedures. Then, IHC and ISH were performed automatically with the Leica Bond-MAX staining system (Leica Biosystems, Hamburg, Germany) by applying mouse monoclonal antibodies raised against LMP-1 (M0897, Dako, Copenhagen, Denmark), EBNA-2 (10R1655, Fitzgerald, Acton, U.S.A.) and ZEBRA (sc-53904, Santa Cruz Biotechnology, Dallas, U.S.A), as well as a fluorescein-conjugated oligonucleotide EBER probe (PB0589, Leica Biosystems) [38, 39].

Chromosomal preparation and in situ hybridisation

Chromosomes were prepared according to standard procedures with an air-drying method [40]. FISH was done as previously reported [41]. As probes, all human whole chromosome paints (wcps) combined in a so-called multicolour-FISH probe set were labelled, applied, processed, and evaluated as previously described [42]. Virus DNA was labelled with SpectrumGreen by nick translation [43] and used in a single-colour FISH assay and/or together with the multicolour-FISH probeset using wcps. Also, nick translation was used to label the probes from MZ1 to 4 in different fluorophores and to simultaneously use in FISH.

Results

Almost the complete HVMA genome sequence was determined

Total DNA of MAL-1 cells was extracted and analysed in a sequencing run using the Illumina HiSeq2000 platform [44]. De novo assembly yielded 19 contigs representing 125,541 nucleotides (nt) of the genome (75.4%). Genome regions corresponding to the larger sequence gaps were amplified, cloned into a plasmid, and the plasmids were sequenced using an Illumina MiSeq run. De novo assembly yielded 22 contigs that represented 62,534 nt (37.5%). The smaller gaps were sequenced by conventional Sanger sequencing. The assembled genome of HVMA consists of 166,590 bp but two DNA stretches within the BHLF1 gene—each of 600–900 bp—remained unsequenced. Circa 93% of the genome was unveiled by next-generation sequencing (NGS) technology while approx. 5.6% was determined by Sanger sequencing. Comparison of HVMA sequence and sequences of EBV and McHV-4 revealed a similar genome organisation of HVMA and EBV. HVMA contains unique and repetitive sequences, including four internal (IR1, IR2, IR3 and IR4) as well as numerous direct repeats (Fig. 1). The 3077-bp unit of the major internal repeat (IR1) has 84.4% nucleotide identity to the 3072 bp unit of EBV-IR1 (HHV-4 wild-type strain B95-8, GenBank accession numbers: AJ507799, NC_007605). The copy number of IR1 repeat units was inferred by comparing the read coverage of the IR1 repeat and those of the remaining genome (single copy) obtained by mapping the Illumina reads to the HVMA genome in which the IR1 repeats were collapsed to a single repeat unit. The presence of five complete and one

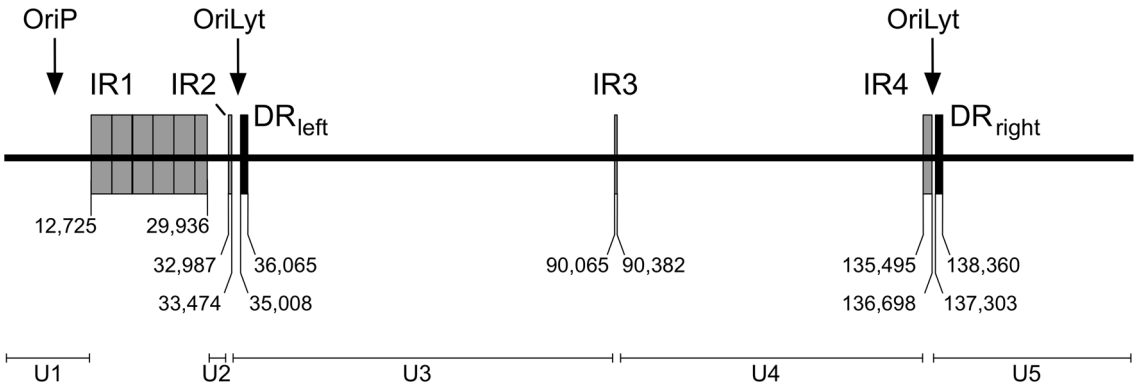

Fig. 1 Linear MarcGHV-1 (HVMA) genome. The assembled genome consists of 166,590 bp. The relative position of internal repeats (IR1-4) as well as direct repeats (DR), the origins of replication (ori) and the unique regions (U) is given. Available data indicate five complete (3077 bp each) and one truncated repeat units of the IR1. The IR2 comprises of 122 bp and may contain of up to 11 repeats. The EBNA-1 open reading frame (ORF) contains IR3-encoded glycine–alanine–serine repeats. The IR4 is part of the LF3 ORF and contains one of the two origins of lytic replication (oriLyt). Two direct repeats (DR) with a length of 1058 bp each were identified downstream of IR2 and IR4. A terminal repeat sequence was not found

partial copies is concluded. The last repeat is truncated and comprises 1827 bp which corresponds to 59.4% of the IR1 unit length. The unit length of IR2 is 122 bp. Four repeats could be identified by Sanger sequencing. According to the size of the underlying PCR product, this genome part was estimated to span approx. 1350 bp, which allows 11 repeats. The EBNA-1 open reading frame (ORF) contains IR3 which encodes an aa repeat [45, 46]. Whereas EBV codes for hundreds of glycine and alanine residues, the HVMA sequence encodes glycine–alanine–serine repeats. The fourth internal repeat, IR4, is part of the LF3 ORF, which also contains one of the two origins of lytic replication (oriLyt). Furthermore, two direct repeats (DR) with a length of 1058 bp each, were identified downstream of IR2 and IR4 (nt 35,008–36,065 and nt 137,303–138,360). A terminal repeat sequence [47] was not found. The sequence was deposited in the GenBank (Acc. No. MG471437).

The HVMA genome organisation was found to be similar to that of EBV

The HVMA genome harbours 91 ORFs. Eighty-six ORFs are unique while one ORF is present five times (BCRF2, BWRF1.2–BWRF1.5). Several attempts failed to determine by conventional Sanger sequencing the complete sequence of the BHLF1 gene which overlaps in part with the OriLyt. Each ORF has a correlate in the EBV genome and a corresponding localisation within the HVMA genome. Therefore, we adopted the designations of the EBV ORFs for HVMA (Fig. 2; Table 1). The highest degree of similarity between HVMA and EBV (GenBank AJ507799, NC_007605) based on the aa sequence was found in the BDLF1 gene encoding a capsid protein (97.4%). Amino acid identities below 50% were observed for the latency proteins EBNA-2, EBNA-3A, -3B, -3C, LMP-1 as well as the LF3. The remaining latency proteins showed medium similarities to their EBV homologues (EBNA-1: 66.1%, EBNA-LP: 57.4%, LMP-2A: 55.0% and LMP-2B: 58.7%).

Phylogenetic analyses support assignment of HVMA to the lymphocryptoviruses of Old World monkeys

To demonstrate close relationship of HVMA with other members of the genus *Lymphocryptovirus*, the DNA polymerase (Pol) aa sequences of 100 gammaherpesviruses were aligned with five reference sequences representing the subfamilies *Alpha-* and *Betaherpesvirinae*. This alignment was used for phylogenetic analysis applying the maximum likelihood method. For ML tree inference the Le_Gascuel_2008 model [48] and a discrete gamma distribution was assumed according to model test. The resultant tree revealed clades consistent with the presently known herpesvirus species [11, 49]. HVMA was found to be most closely related to lymphocryptoviruses of Old World monkeys and EBV, but distant from New World monkey lymphocryptoviruses (Fig. 3).

Previously, it was demonstrated that all members of the *Herpesviridae* share so-called "core genes" presumably derived from a common ancestor [36]. In this study, 42 HVMA core genes (71,110 nt) were compared to the corresponding concatenated genes of 21 members of *Gammaherpesvirinae*, including all four genera, as well as four species of the *Alpha-* and *Betaherpesvirinae* (see Table S1 for details). The ML tree was calculated applying the general time reversible substitution model [50], a discrete gamma distribution and invariable rate variation (GTR + G + I). Again, the HVMA sequence clusters with those of lymphocryptoviruses of Old World monkeys (Fig. 4). According

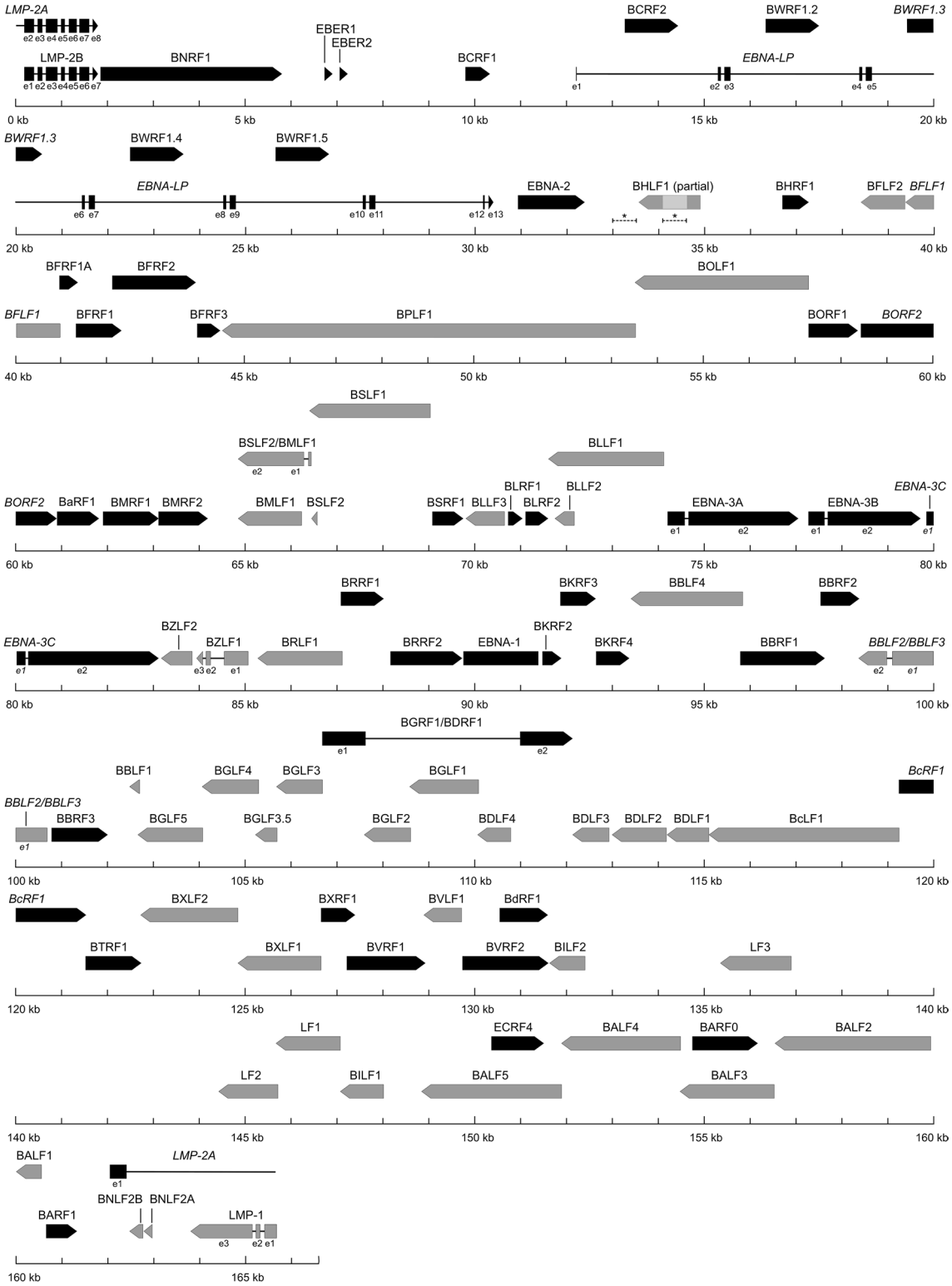

Fig. 2 Organisation of the MarcGHV-1 (HVMA) genome. Indicated are 86 unique open reading frames (ORFs) and non-coding EBER-like genes. The latent genes are named after their proteins, the lytic genes in analogy to the nomenclature of the EBV-BamHI-reading frames. One ORF is present five times (BCRF2, BWRF1.2–BWRF1.5). The BHLF1 ORF is incomplete, two regions remained unsequenced (*). ORF sizes are drawn to scale. Orientation of ORFs and EBER RNAs is indicated by direction of the arrow and shading (black = rightward, grey = leftward). Italic font indicates line break of ORF

Table 1 Predicted MarcGHV-1 (HVMA) genes and gene products compared to EBV

MarcGHV-1					EBV		
Gene	Exon	Start (nt)	Stop (nt)	Size (aa)	Size (aa)	% aa identity	Function
LMP-2A		161,746	1778	489	498	55.0	Important for latency
	LMP-2A e1	162,047	162,402				
	LMP-2A e2	179	393				
	LMP-2A e3	473	571				
	LMP-2A e4	652	900				
	LMP-2A e5	981	1065				
	LMP-2A e6	1145	1319				
	LMP-2A e7	1376	1591				
	LMP-2A e8	1669	1778				
LMP-2B		180	1778	370	378	58.7	Important for latency
	LMP-2B e1	180	393				
	LMP-2B e2	473	571				
	LMP-2B e3	652	900				
	LMP-2B e4	981	1065				
	LMP-2B e5	1145	1319				
	LMP-2B e6	1376	1591				
	LMP-2B e7	1669	1778				
BNRF1		1835	5788	1317	1318	79.4	Tegument protein
EBER1		6722	6895				Small RNA
EBER2		7052	7219				Small RNA
BCRF1		9786	10,319	177	176	84.1	IL-10 homologue
EBNA-LP		12,198	30,298	365	506	57.4	Nuclear protein
	EBNA-LP e1	12,198	12,199				
	EBNA-LP e2	15,289	15,349				
	EBNA-LP e3	15,431	15,562				
	EBNA-LP e4	18,366	18,426				
	EBNA-LP e5	18,508	18,639				
	EBNA-LP e6	21,443	21,503				
	EBNA-LP e7	21,585	21,716				
	EBNA-LP e8	24,520	24,580				
	EBNA-LP e9	24,662	24,973				
	EBNA-LP e10	27,597	27,657				
	EBNA-LP e11	27,739	27,870				
	EBNA-LP e12	30,081	30,113				
	EBNA-LP e13	30,197	30,298				
BCRF2 (BWRF1.1)		13,257	14,414	385	383	70.3	
BWRF1.2		16,334	17,491	385	383	70.3	
BWRF1.3		19,411	20,568	385	383	70.3	
BWRF1.4		22,488	23,645	385	383	70.3	
BWRF1.5		25,565	26,722	385	383	70.3	
EBNA-2		30,843	32,291	482	487	46.1	Nuclear protein
BHLF1 (partial)		34,912	33,572	>447	660	–	Phosphoprotein 85
BHRF1		36,699	37,271	190	191	68.0	bcl-2 homologue
BFLF2[a]		39,368	38,412	318	318	84.6	Nuclear egress lamina protein
BFLF1[a]		40,955	39,381	524	525	86.0	Nuclear protein
BFRF1A[a]		40,954	41,355	133	135	86.9	Terminase associated
BFRF1[a]		41,315	42,310	331	336	80.1	Nuclear protein
BFRF2		42,214	44,034	606	591	82.0	Nuclear egress membrane protein

Table 1 (continued)

MarcGHV-1					EBV		
Gene	Exon	Start (nt)	Stop (nt)	Size (aa)	Size (aa)	% aa identity	Function
BFRF3[a]		43,958	44,464	168	176	77.3	Small capsid protein
BPLF1[a]		53,502	44,506	2998	3149	76.2	Tegument protein
BOLF1[a]		57,274	53,489	1261	1239	78.2	Interacts with tegument protein (BPFL1)
BORF1[a]		57,273	58,346	357	364	91.3	Capsid structure
BORF2[a]		58,409	60,886	825	826	88.5	Ribonucleotide reductase
BaRF1[a]		60,899	61,807	302	302	95.1	Ribonucleotide reductase
BMRF1[a]		61,900	63,114	404	404	89.5	DNA polymerase complex
BMRF2		63,119	64,195	358	357	57.0	Membrane protein
BMLF1		66,216	64,837			83.1	
BSLF2		66,314	66,436	40	39	66.7	
BSLF2/BMLF1[a]		66,436	64,837	498	479	81.9	mRNA export
	BSLF2/BMLF1 e2	66,269	64,837				
	BSLF2/BMLF1 e1	66,436	66,373				
BSLF1[a]		69,032	66,405	875	874	87.5	Helicase–primase subunit
BSRF1		69,075	69,746	223	218	90.4	Tegument protein
BLLF3[a]		70,643	69,804	279	278	89.0	dUTPase
BLRF1[a]		70,717	71,025	102	102	73.9	Glycoprotein N
BLRF2		71,096	71,587	163	162	77.6	Capsid protein
BLLF1		74,111	71,604	835	907	64.3	Glycoprotein gp350
BLLF2		72,163	71,744	139	148	61.4	Function unknown
EBNA-3A		74,202	77,044	919	944	41.7	Nuclear protein
	EBNA-3A e1	74,202	74,567				
	EBNA-3A e2	74,651	77,044				
EBNA-3B		77,271	79,723	791	938	48.3	Nuclear protein
	EBNA-3B e1	77,271	77,621				
	EBNA-3B e2	77,699	79,723				
EBNA-3C		79,848	83,120	1064	992	46.3	Nuclear protein
	EBNA-3C e1	79,848	80,198				
	EBNA-3C e2	80,277	83,120				
BZLF2		83,841	83,167	224	223	80.7	Glycoprotein gp42
BZLF1		85,065	83,940	248	245	74.5	ZEBRA protein
	BZLF1 e3	84,065	83,940				
	BZLF1 e2	84,259	84,155				
	BZLF1 e1	85,068	84,553				
BRLF1		87,117	85,282	611	605	79.5	Transactivator
BRRF1		87,095	88,024	309	310	88.3	Transcription factor
BRRF2		88,171	89,730	519	537	64.3	Tegument protein
EBNA-1		89,765	91,396	543	641	66.1	Nuclear protein
BKRF2[a]		91,478	91,891	137	137	83.2	Glycoprotein L
BKRF3[a]		91,873	92,640	255	255	95.1	Uracil-DNA glycosylase
BKRF4		92,651	93,370	239	217	68.5	Tegument protein
BBLF4[a]		95,844	93,412	810	809	93.4	Helicase–primase complex
BBRF1[a]		95,789	97,633	614	613	93.2	Virion protein
BBRF2[a]		97,536	98,372	278	278	92.3	Tegument protein
BBLF2/BBLF3[a]		100,684	98,369	721	723	77.9	Helicase–primase
	BBLF2/BBLF3 e2	98,982	98,369				
	BBLF2/BBLF3 e1	100,684	99,133				
BBRF3[a]		100,784	102,004	406	405	92.1	Glycoprotein M

Table 1 (continued)

MarcGHV-1					EBV		
Gene	Exon	Start (nt)	Stop (nt)	Size (aa)	Size (aa)	% aa identity	Function
BBLF1[a]		102,706	102,488	72	75	70.8	Tegument protein
BGLF5[a]		104,073	102,661	470	470	93.7	Alkaline exonuclease
BGLF4[a]		105,349	104,057	430	429	93.0	Virion protein kinase
BGLF3.5		105,689	105,225	154	153	87.1	Tegument protein
BGLF3[a]		106,674	105,676	332	332	88.1	Terminase-binding protein
BGRF1/BDRF1[a]		106,673	112,119	690	690	93.8	Terminase
	BGRF1/BDRF1 e1	106,673	107,608				
	BGRF1/BDRF1 e2	110,983	112,119				
BGLF2[a]		108,608	107,598	336	336	88.2	Tegument protein
BGLF1[a]		110,094	108,586	502	507	77.2	Tegument protein
BDLF4		110,786	110,064	240	225	86.7	Function unknown
BDLF3		112,921	112,130	263	234	58.3	Glycoprotein gp150
BDLF2		114,170	112,986	394	420	67.3	Tegument protein
BDLF1[a]		115,085	114,180	301	301	97.4	Capsid structure
BcLF1[a]		119,246	115,101	1381	1381	96.1	Major capsid protein
BcRF1		119,245	121,533	762	750	83.5	Regulation of latent genes
BTRF1[a]		121,520	122,731	403	425	89.2	Cytoplasmatic egress
BXLF2[a]		124,848	122,728	706	706	84.0	Glycoprotein H
BXLF1[a]		126,655	124,850	601	607	90.6	Thymidine kinase
BXRF1[a]		126,654	127,397	247	248	84.6	Function unknown
BVRF1[a]		127,213	128,919	568	570	85.9	Tegument protein
BVLF1		129,704	128,889	271	272	87.9	Function unknown
BVRF2[a]		129,730	131,595	621	605	78.9	Protease
BdRF1[a]		130,543	131,595	350	345	77.5	Capsid protein
BILF2		132,390	131,629	253	248	73.3	Glycoprotein gp78
LF3		137,363	135,354	670	924	46.0	Function unknown
LF2		145,709	144,420	429	429	93.4	Function unknown
LF1		147,064	145,670	464	469	78.6	Function unknown
BILF1		148,011	147,073	312	312	81.8	Glycoprotein gp60
BALF5[a]		151,884	148,837	1015	1015	95.1	DNA polymerase
ECRF4[b]		150,364	151,500	378			Function unknown
BALF4[a]		154,481	151,887	864	857	87.5	Glycoprotein B
BALF3[a]		156,522	154,468	684	789	88.7	Terminase
BARF0		154,741	156,156	471	471	75.2	Function unknown
BALF2[a]		159,928	156,539	1129	1128	94.4	ssDNA-binding protein
BALF1		160,561	160,013	182	220	85.8	bcl-2 homologue
BARF1		160,661	161,326	221	221	75.1	Growth factor
BNLF2B		162,779	162,483	98	101	64.4	Function unknown
BNLF2A		162,967	162,788	59	60	53.8	Immune evasion
LMP-1		165,678	163,814	561	386	49.1	Immortalisation
	LMP-1 e3	165,150	163,814				
	LMP-1 e2	165,321	165,235				
	LMP-1 e1	165,678	165,417				

Position of the genes within the MarcGHV-1 genome and the homology of the translated proteins with EBV (wild type, NC_007605). Homology of the non-translated EBV-encoded RNAs: EBER-1 82.6%, EBER-2 66.6%. Function of EBV proteins according to [78–84]

[a]Proteins conserved within family of *Herpesviridae* (core genes)

[b]ORF in EBV wild-type non-existent

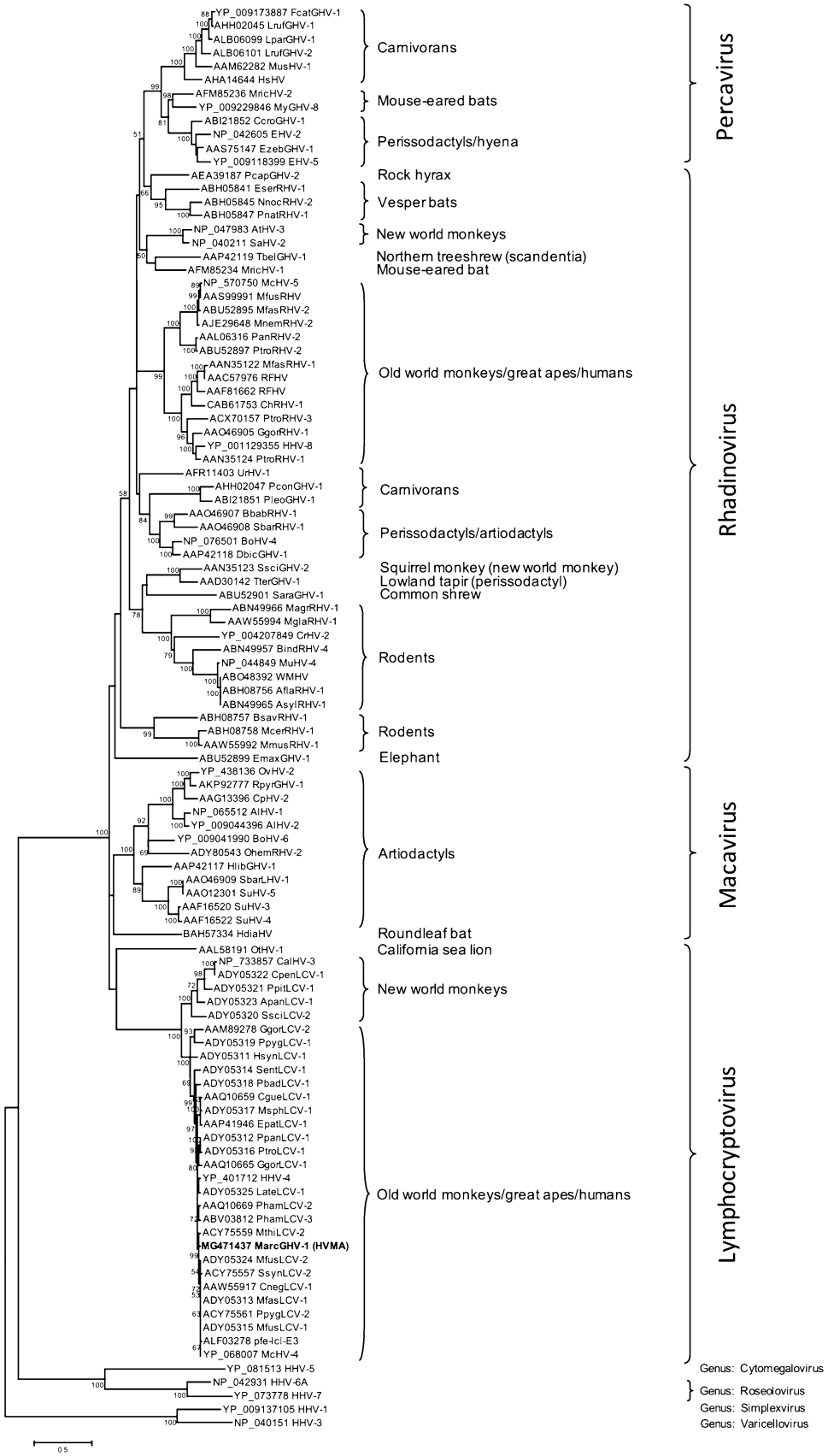

Fig. 3 Phylogenetic analysis of 105 DNA polymerase amino acid sequences. The tree includes sequences of the four gammaherpesvirus genera as well as members of the alpha- and betaherpesvirus subfamilies. Host species are given next to virus abbreviations (for detailed virus names, see Table S3). MarcGHV-1 (HVMA) is printed in bold. Bootstrap support is indicated next to the nodes. The scale bar indicates the number of substitutions per site

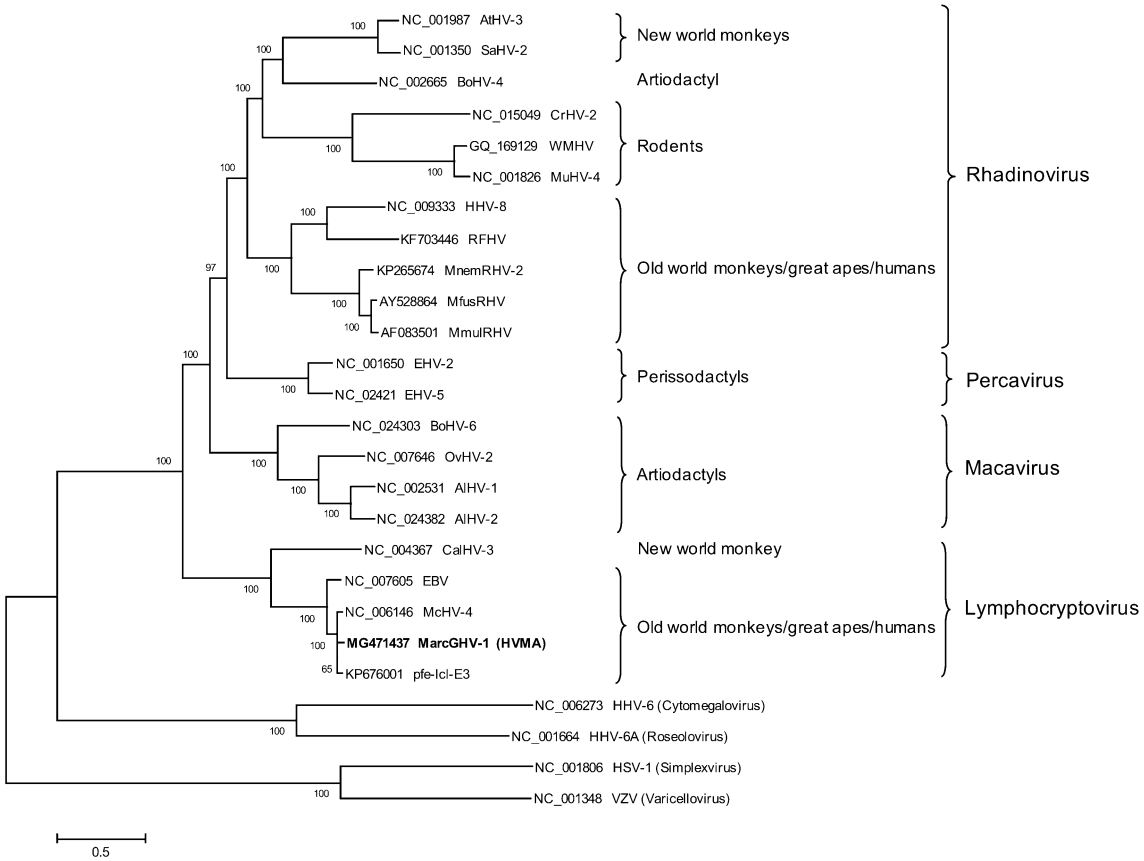

Fig. 4 Phylogenetic analysis of 42 core genes. Twenty-two members of the gammaherpesvirus subfamily as well as two members of the alpha- and the betaherpesvirus subfamilies were included in this tree. The gammaherpesvirus genera are given as well as the host species. Virus abbreviations are listed in Table S3. The MarcGHV-1 (HVMA) sequence is printed in bold. As for the DNA polymerase tree (see Fig. 3), MarcGHV-1 was found to be closely related to other lymphocryptoviruses of Old World monkeys and EBV. The evolutionary history was inferred using the maximum likelihood method based on the general time reversible model. Bootstrap values are presented at the nodes. The scale bar indicates the number of substitutions per site

to the present naming practice of gammaherpesviruses, we suggest the designation 'Macaca arctoides gammaherpesvirus 1' (MarcGHV-1) for this virus.

Presence of episomal and linear MarcGHV-1 (HVMA) DNA was demonstrated in MAL-1 and HTRL

To investigate whether MarcGHV-1 is able to productively replicate in MAL-1 and HTR lymphocytes (HTRL), the presence of episomal covalently closed circular DNA (cccDNA) and linear DNA (that could act as template during early replication steps [51, 52]) was investigated by in situ lysis gels according to Gardella et al. [37]. Raji cells (EBV latency type III) [53] were chosen as a positive control, whereas the EBV-free BL cell line Ramos [27] served as a negative control. Lytic EBV cycle was induced in Raji cells by phorbol ester (12-O-tetradecanoylphorbol-13-acetate, TPA) [28]. A number of 10^6 cells was used for this experiment, which was done in accordance with previous studies [37, 54], followed by Southern blot hybridisation. This technique allows the discrimination of episomal and linear DNA due to their different physical properties [37].

As demonstrated in Fig. 5, a probe specific to BSLF1 and BSRF1 ORFs of lymphocryptoviruses revealed a strong signal in Raji and HTRL cells which corresponds to episomal DNA; signal intensity was weaker in MAL-1 cells. In addition, distinct signals corresponding to linear viral DNA were detected in TPA-treated Raji, MAL-1, and HTRL, but also in untreated MAL-1. A weak smear was present in the corresponding areas of untreated HTRL and Raji cells. The signals found in the wells of the gel should represent viral DNA being integrated into the cellular genome [37], but also some unspecific probe binding to genomic DNA. Thus, this reactivity was also observed in the well containing virus-free Ramos cells and was not taken into consideration.

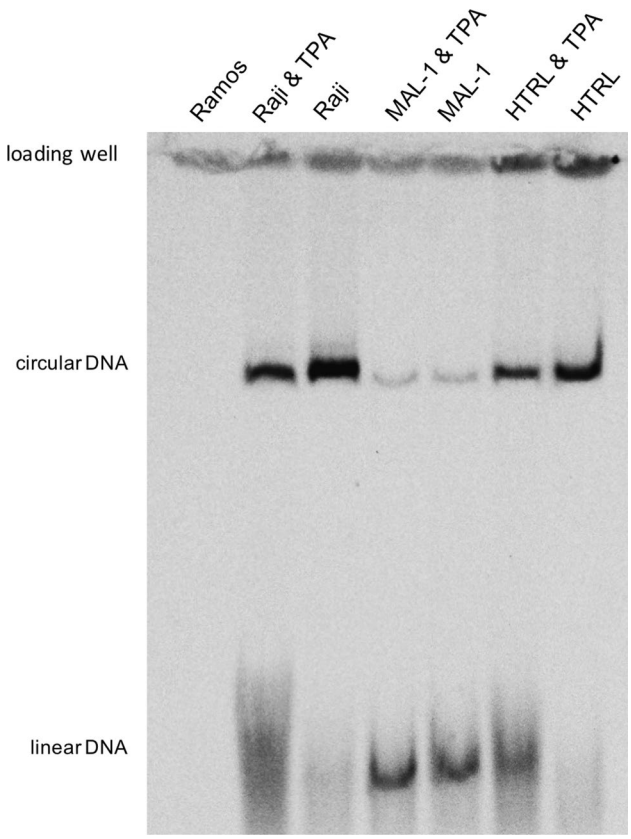

Fig. 5 Southern blot hybridisation of in situ lysis gel. Approximately 1×10^6 cells were loaded per slot and fractionated by electrophoresis. Then, DNA was detected by Southern blot analysis using a gammaherpesvirus-specific probe. Latently EBV-infected Raji cells (50–60 copies of episomal DNA per cell) served as positive control, and EBV-negative Ramos cells as mock control. The lytic replication cycle was chemically induced with TPA 48 h before electrophoresis. Circular DNA was detected in all virus-infected cells. However, slots loaded with MAL-1 cells exhibit weaker signal intensities. Linear DNA, which indicates active MarcGHV-1 (HVMA) replication, could be detected in the slots carrying MAL-1 cells and TPA-treated HTRL and Raji cells

Proviral DNA of STLV was not found in MAL-1 and HTRL

Nucleic acid preparations of both cell lines, MAL-1 and HTRL, were tested by nested PCR for the presence of proviral STLV-1 *tax* gene DNA [30]. A DNA preparation containing approx. 300 copies of proviral HTLV-1 DNA, which is very closely related to STLV-1 [55], was used as positive control. Applying this highly sensitive method, neither MAL-1 nor HTRL were found to harbour STLV-1 proviral DNA (Fig. S1). Retroviruses, however, may gather nucleotide substitutions after insertion into the host cell genome [56]. These could act as confounders of PCR amplification. Thus, we also tried to map the Illumina data of the MAL-1 sequencing run (appr. 100 mio reads, average coverage 6.45) to a set of 377 STLV-1 and STLV-3 sequences available in GenBank (data as of – 1 February 2018). No significant evidence of STLV sequences was found in the MAL-1 sequence dataset which is consistent with the assumption that transformation of MAL-1 cells was not caused by integration of a STLV-like retrovirus [22].

B- and T-cell-specific transcripts were evident in MAL-1 and HTRL

To determine the cell type of MAL-1 and HTRL, the expression of specific markers was investigated by RNA sequencing (Illumina) using total RNA of TPA-treated and untreated cells. The Illumina reads were mapped to the available transcript sequences (data as of 4 April 2018) of cellular markers for B-lymphocytes (CD19, CD20, CD21 and CD79α), T-lymphocytes (CD2, CD3ε/γ, CD4, CD8α/β), but also for granulocytes, monocytes and macrophages (CD14, CD93, CD163) of *Macaca arctoides, Macaca fascicularis, Macaca mulatta, Macaca nemestrina* and *Oryctolagus cuniculus*, respectively. Cell-type specificity was assumed in analogy to cluster of differentiation (CD) nomenclature from human [57]. The B-cell-specific transcript for CD79α and the T-cell-specific transcripts CD3ε/γ were the most abundant marker RNAs found in both cell lines. The HTRL RNA preparation was also positive for CD4 and CD8 (Fig. 6a, b; Table 2).

Slight differences in the transcription and expression patterns of EBV-latency homologs were found between MAL-1 and HTRL

Three latency types have been described for EBV-transformed cells: type I latency is characterised by expression of EBV nuclear antigen (EBNA)-1 only, type II latency by expression of EBNA-1, latent membrane protein (LMP)-1, -2A and -2B, whereas in type III all nine latency proteins are expressed (EBNA-1, -2, -3A, -3B, -3C, -LP, LMP-1, -2A, -2B). In addition, EBV-transformed cells transcribe two small EBV-encoded RNAs (EBER1, 2) [58]. Two independent methods were used to investigate the type of viral latency in TPA-treated and native MAL-1 and HTRL, and results are summarised in Table 3. First, RNA was tested for the presence of MarcGHV-1-specific transcripts encoding EBNA-1, -2 and -3A–C, the LMP-1 and -2A/B, the lytic switch transactivator ZEBRA and the early genes BSLF2/BMLF1 by Illumina sequencing technology. Thereby, the presence of EBV-encoded RNA homologs to EBER1 and more abundantly to EBER2 was demonstrated in MAL-1 and HTRL. The EBNA-1-specific transcript was also present in both cell lines while transcription of EBNA-2 was observed only in MAL-1. Differences in the transcription pattern were also demonstrated for EBNA-3A, -3B, and -3C, all of which

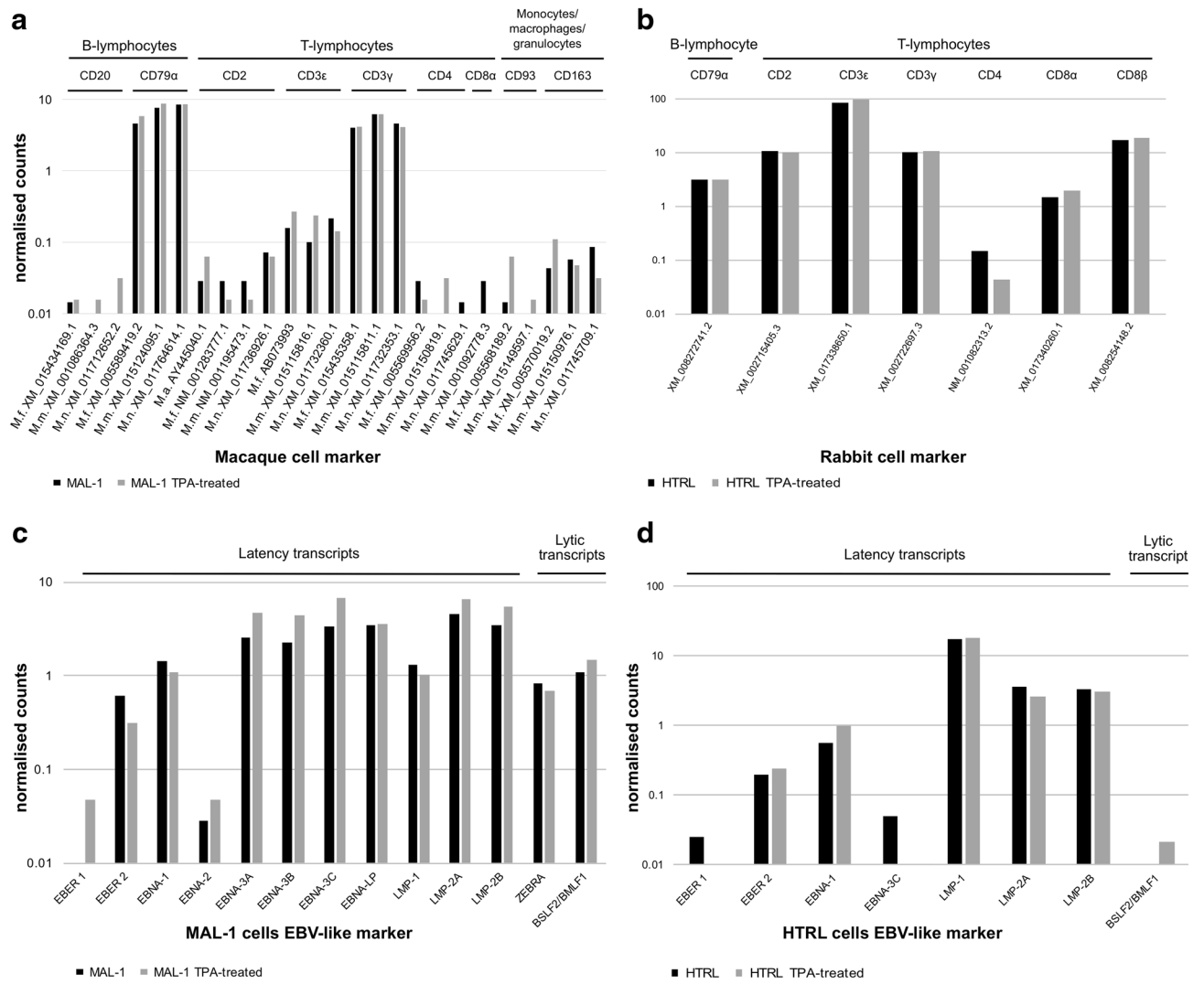

Fig. 6 Determination of cell type and presence of EBV-like latency transcript homologs using transcription analysis. Complementary DNA libraries from RNA preparations of TPA-treated and untreated MAL-1 cells (**a, c**) and HTRL (**b, d**) were sequenced with the Illumina HiSeq2500 platform. **a, b** Resulting reads were mapped to transcript sequences from GenBank (data as of 4 April 2018) specific for B-lymphocytes (CD19, CD20, CD21 and CD79α) and T-lymphocytes (CD2, CD3ε/γ, CD4, CD8α/β) as well as granulocytes/monocytes/macrophages (CD14, CD93, CD163) of macaques and rabbits. No reads mapping to CD19, CD21, CD8β, and CD14 were found in MAL-1 cells, whereas no reads mapping to CD19, CD20, CD21, CD14, CD93, and CD163 were detected in HTRL. As different transcript variants of cell markers are available, only the reads with the highest signal are shown. **c, d** Resulting reads were mapped to EBV-like latency (EBER1/2, EBNA-1/-2/-3A/-3B/-3C, LMP-1/-2A/-2B) as well as lytic markers ZEBRA (BZFL1) and BSLF2/BMLF1 of MarcGHV-1. No reads mapping to EBNA-2, -3A, -3B, and ZEBRA were found in HTRL. *M.a. Macaca arctoides, M.f. Macaca fascicularis, M.m. Macaca mulatta, M.n. Macaca nemestrina*

Table 2 Cell marker to determine origin of MarcGHV-1 (HVMA) transformed cells

	MAL-1	MAL-1 TPA treated	HTRL	HTRL TPA treated
B-lymphocytes	CD79α	CD79α	CD79α	CD79α
T-lymphocytes	CD2, CD3ε, CD3γ	CD2, CD3ε, CD3γ	CD2, CD3ε, CD3γ, CD4, CD8α, CD8β	CD2, CD3ε, CD3γ, CD4, CD8α, CD8β

Summarised data from Illumina sequencing (compare Fig. 6a, b)

Table 3 EBV latency and lytic markers present in MarcGHV-1 (HVMA) transformed cells

	MAL-1	MAL-1 TPA treated	HTRL	HTRL TPA treated
Latent infection				
EBER	+	+	+	+
EBNA-1	+	+	+	+
EBNA-2	+	+	–	–
EBNA-3A	+	+	–	–
EBNA-3B	+	+	–	–
EBNA-3C	+	+	+	–
LMP-1	+	+	+	+
LMP-2A	+	+	+	+
LMP-2B	+	+	+	+
Lytic infection				
ZEBRA	+	+	–	–
BSLF2/BMLF1	+	+	–	+

Summarised data from Illumina sequencing and in situ hybridisation/immunohistochemistry (compare Figs. 6c, d, 7a, b)

were present in MAL-1, while only EBNA-3C was seen in HTRL. The transcripts of LMP-1 and LMP-2A/B homologs were demonstrated in both cell lines. In addition, transcripts for ZEBRA and BSLF2/BMLF1 homologs were present in MAL-1. Treatment with TPA had only a minor effect on the expression levels of transcripts (Fig. 6c, d).

Second, formalin-fixed and paraffin-embedded cells were analysed for the presence of EBER RNAs and the expression of EBNA-2, LMP-1 and ZEBRA by ISH/IHC (Fig. 7). TPA-treated and -untreated Raji cells served as positive control while Ramos cells were used as negative control (Fig. S2). The MAL-1 cells exhibited EBER RNAs in their nucleus as well as EBNA-2. This is confirmed by the Illumina sequencing results. Few cells were also tested positive for ZEBRA expression in the nucleus and cytoplasm. Weak LMP-1 expression was only seen after TPA treatment (Fig. 7a). HTRL, however, showed only EBER signal, irrespective of TPA treatment (Fig. 7b), and expression of EBNA-2, LMP-1 and ZEBRA was not demonstrated. This is not confirmed with the results of the transcription analysis and may reflect lower protein levels.

Chromosomal aberrations were evident in MAL-1

Cells of MAL-1 and HTRL were cytogenetically worked up, fixed on slides and metaphase chromosomes were analysed using inverted DAPI banding technique. In MAL-1 a karyotype typically seen in macaque species was identified (data not shown). As previous studies have demonstrated the suitability of a human multicolour-FISH probe set consisting of all human whole chromosome paints (wcps) [59–61] for cytogenetic analyses of macaque cells, this probe set was also applied to MAL-1 cells. The MAL-1 cells appeared to have a stable, rearranged, nearly-tetraploid macaque karyotype. Few metaphases presented some non-clonal additional aberrations. The MAL-1 cell karyotype could be defined as follows:

79,XXXX,der(1),der(1),der(3),der(3)t(2;3)x2,-5,der(5)t(X;5)x2,der(5)t(2;5;17),der(7)t(3;7),-8,-8,der(8)t(8;16)x2,der(12)t(8;12),der(12)t(8;12),der(14)t(1;14),-15,-17,der(17)t(2;17)x2,der(18)t(2;18)x2 (Fig. S3).

HTRL cells were also karyotyped, and based on the inverted DAPI banding pattern it was identical to the published rabbit karyotype [62] (Fig. S4). The different host species origin of high-passage lymphocyte preparation was re-confirmed by species-specific amplification of mitochondrial DNA according to Ono et al. [63] (data not shown) as well as by detection of rabbit specific transcripts (see above).

Insertion of MarcGHV-1 DNA into the host genome was demonstrated by FISH on MAL-1 and HTRL chromosomal preparations

Four probes (MZ1-4), representing different parts of the MarcGHV-1 genome, were generated from cloned PCR products. The probes were labelled with biotin (MZ1, detected by streptavidin-FITC or streptavidin-Cy5), SpectrumOrange (MZ2), SpectrumGreen (MZ3) and DY-415 (MZ4), respectively.

The MarcGHV-1 genome specificity was exemplarily shown for the MZ1 probe, which exhibited no signal in human and macaque control metaphases as well as in EBV-transformed human cell lines 13L0259, 13L0260 and 13L0158 from the Else Kröner-Fresenius-sSMC-cell bank (http://ssmc-tl.com/ekf-cellbank.html) (data not shown). In contrast, multiple inserts were seen in MAL-1 chromosomal preparations. The number of MZ1 signals and signal intensity varied in different metaphases, which made determination of the number of insert copies per cell impossible by this technique. However, it was assumed that large bright signals indicate insert amplifications, whereas small single copy inserts are barely visible (Fig. 8a). Multiple MarcGHV-1 inserts were observed in HTRL metaphases in almost in every chromosome. In general, the number and intensity of signals appeared to be higher in HTRL than in MAL-1 cells (Fig. 8b). In the next step, all four probes were mixed and applied to MAL-1 metaphases. Signal co-localisation was observed, most likely due to chromosomal integration of complete MarcGHV-1 genomes (Fig. 8c, d).

Then, MAL-1 chromosomes with integrated MarcGHV-1 DNA were identified, applying the MZ1 probe and a human multicolour-FISH probe mix based on wcps in parallel on 20 metaphases. All metaphases showed inserts, but the number of inserts per cell and the affected chromosomes varied. The

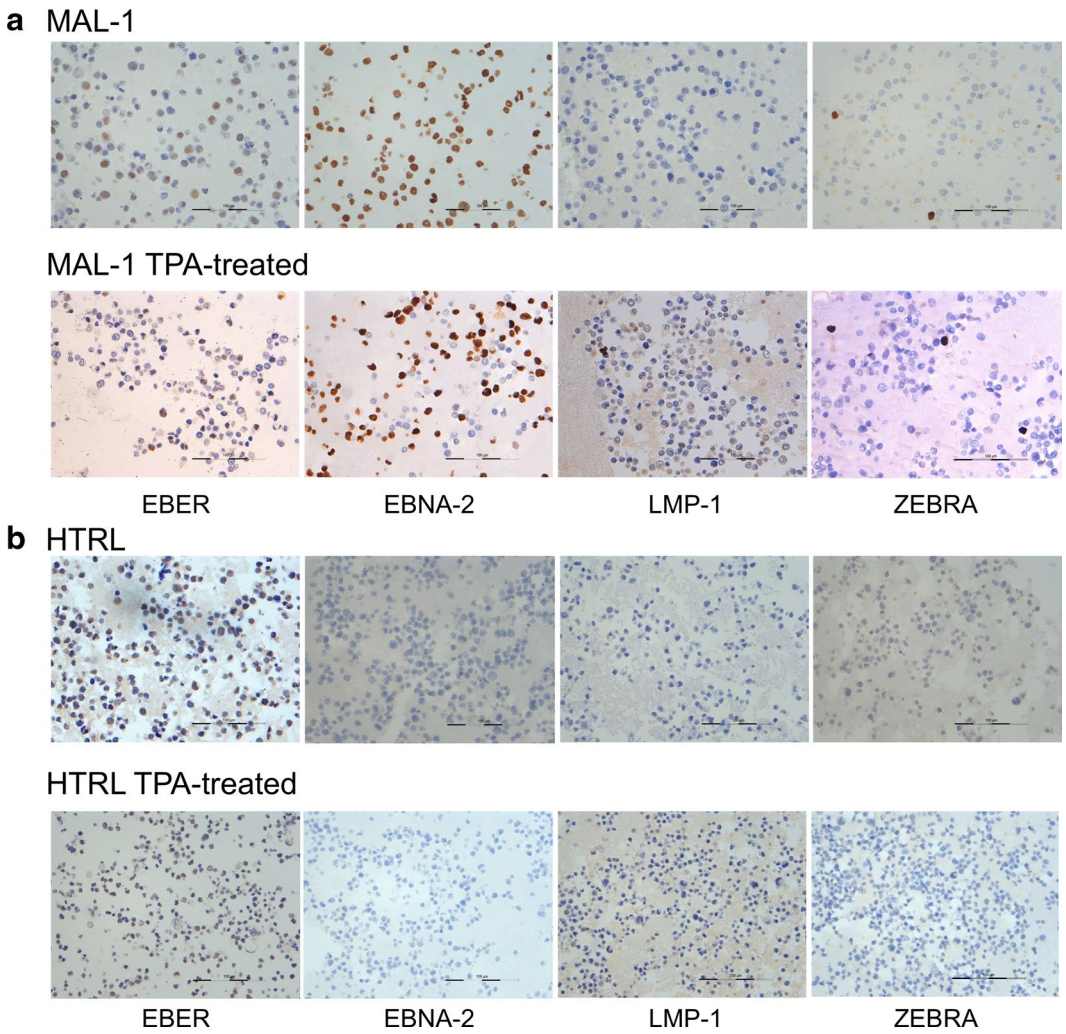

Fig. 7 Detection of EBER by in situ hybridisation and antigenic profile of native and TPA-treated cells. MAL-1 cells (**a**) and HTRL (**b**) grown in suspension were pelleted by centrifugation. After several washings and formalin treatment, the remaining pellet was paraffin embedded. EBER detection and immunohistochemistry using monoclonal antibodies against EBV proteins EBNA-2, LMP-1 and ZEBRA was performed on 2-µM sections. EBV-positive Raji and EBV-negative Ramos cells served as a control (compare Fig. S2). Size bar = 100 µm

chromosomes #4 (17 of 20 metaphases), #6 (16/20) and #16 (12/20) were mostly affected by MarcGHV-1, with a similar localisation of DNA integration, respectively (Table S2 and Fig. S5). The application of several commercially available FISH probes together with MZ1-probe provided no evidence for the existence of chromosomal translocations at the points of and in connection with virus integration sites in MAL-1 cell metaphases (data not shown).

Illumina sequencing data confirmed chromosomal integration of MarcGHV-1 DNA

To analyse whether the MarcGHV-1 genome integrates into the host genome and to identify the site of integration, the MAL-1 read pairs were mapped to the MarcGHV-1 genome first, followed by mapping the not-mappable read mates of the unpaired mappable read pairs to the genome of *Macaca mulatta* (rheMac8). This resulted in 414 read pairs with one mate mapping to the MarcGHV-1 genome and the corresponding mate mapping to the macaque genome. Then, the reciprocal approach was done using another macaque genome (*Macaca fascicularis*). First, the read pairs were mapped to the macaque genome (macFas5), followed by mapping to the MarcGHV-1 genome. This approach resulted in 418 read pairs. The intersection of both approaches included 404 read pairs. Analysis of these overlapping read pairs revealed numerous integration sites of the host genome. This is consistent with the results of the FISH experiment which suggested a polyclonal tumour and integration in various sites of various

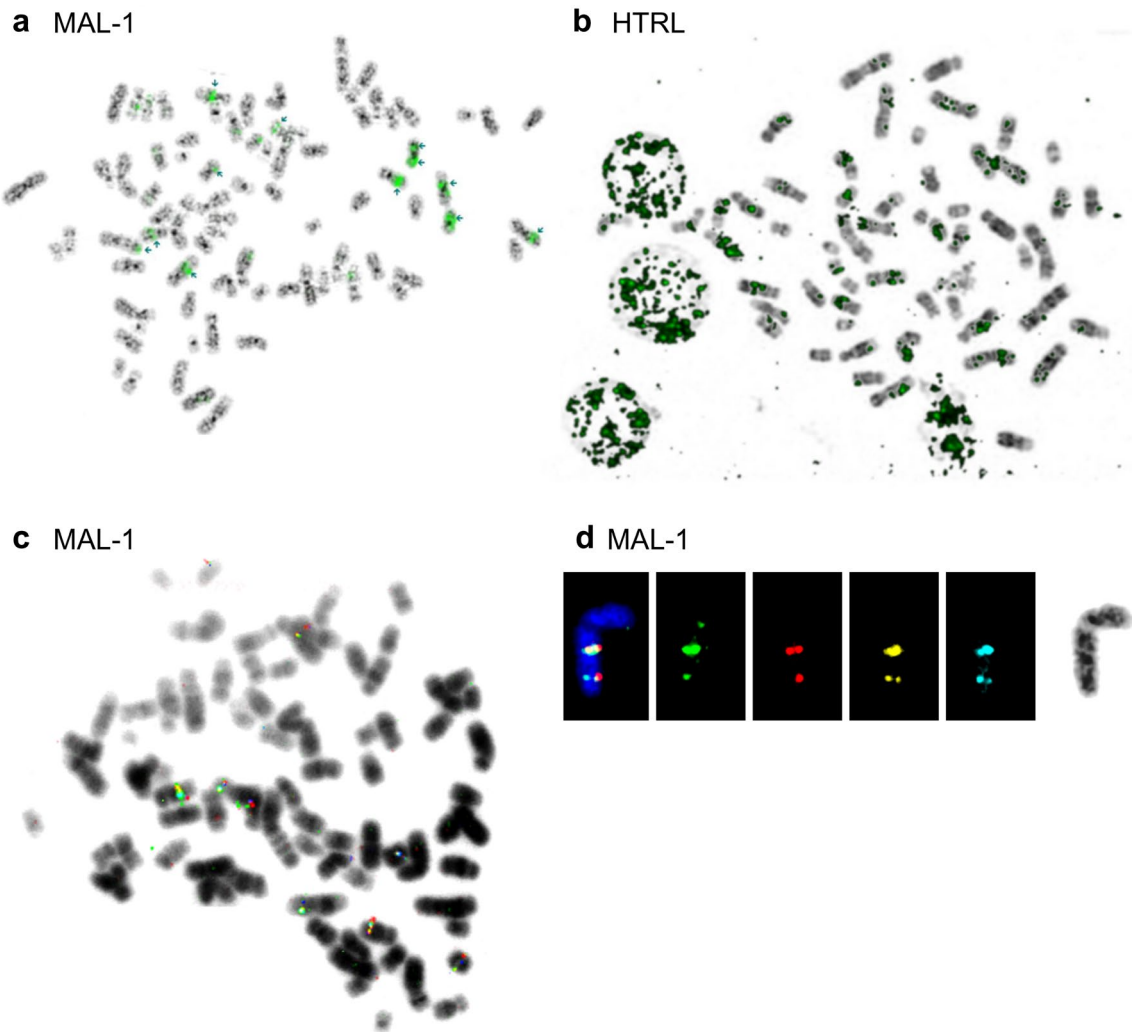

Fig. 8 Fluorescence in situ hybridisation of a MarcGHV-1-derived probe on metaphases. Merged image of inverted DAPI and MZ1 signals (green; = MarcGHV-1 derived probe) on MAL-1 cell (**a**) and HTRL (**b**) metaphase. The distribution of the MarcGHV-1 DNA over the whole genome is clearly visible in both cell types; besides three interphase nuclei full of specific signals for MarcGHV-1 derived probe are visible in **b**. **c** Merged image of inverted DAPI and MZ probe mix representing four parts of the viral genome: MZ1 (Bio/SA-Cy5), MZ2 (Spectrum Orange), MZ3 (Spectrum Green), and MZ4 (DY-415) on MAL-1 cell metaphase. **d** Representative fluorochrome profile of one of the chromosomes from MAL-1 cell metaphase hybridised with MarcGHV-1-derived probe mix. Left to right—merged image, MZ3 (Spectrum Green), MZ2 (Spectrum Orange), MZ1 (Biotin-Straptavidin-Cy5), and MZ4 (DY-415), inverted DAPI image. Overall, this shows the co-localisation of all four used probes

chromosomes. In addition, no preferred sites of recombination were detected in the virus genome.

Discussion

The MarcGHV-1-carrying cell line MAL-1 was established from peripheral lymphocytes of *Macaca arctoides* [18]. This virus was able to transform rabbit lymphocytes [19] and induced malignant lymphomas in rabbits, although there were some controversies about a possible contribution of STLV-1 in this matter [22–24]. The latter may have limited the proposed suitability of MarcGHV-1-infected macaques and rabbits as animal models for EBV-like lymphoproliferative disease [19–23]. We reasoned that a better characterisation of MarcGHV-1 and its host cells may facilitate the availability of such models; which is, so far, limited to humanised mice [64–66].

Here, the genome of the HVMA strain of MarcGHV-1 was sequenced by a combination of NGS- and Sanger technology. Similar to other gammaherpesviruses, HVMA contains unique and repeated sequences, including four internal repeats and two direct repeats (Fig. 1). The size of the HVMA genome is 167 kbp and fits well with a

previous estimation of approximately 170 kbp which was obtained using restriction enzyme-digested DNA and blotting techniques [67]. HVMA and EBV have a very similar genome organisation (Fig. 2; Table 1). Sequence gaps that could not be closed remain in the BHLF1 gene which overlaps with the IR2. Approximately 600 nt (200 amino acids) at the 5′-end and 860–900 nt of the 3′-end including 7 of 11 IR2 repeat units and the BHLF1 polyadenylation signal are still missing. This equals roughly one per cent of the whole virus genome.

Sequence interpretation has to take into account that the consensus sequence of the de novo generated HVMA genome resulted from total DNA preparations of the transformed MAL-1 cell line. Thus, a mixture of defective, integrated viral sequences and intact, episomal genomes cannot be excluded *a priori*. The almost identical genome organisation compared to EBV, however, suggests that this is not a major problem for the re-construction of a consensus sequence. Nevertheless, particularly the estimation of repeat numbers still contains some uncertainty, which cannot be solved by the Illumina sequencing method used. Whether third-generation methods, which generate long and very long reads, are an useful alternative, has to be investigated in further studies. The observation that MAL-1 cells constitute a polyclonal tumour with many different integration sites could confound this approach.

Phylogenetically, HVMA is closely related to other lymphocryptoviruses of Old World monkeys and EBV, but distant from New World monkey lymphocryptoviruses (Figs. 3, 4). With respect to DNA polymerase aa sequence data, our phylogenetic analysis failed to resolve perca- and rhadinovirus species which may be addressed in further studies (Fig. 3).

Using in situ lysis gel electrophoresis technique [37], the presence of episomal, covalently closed circular MarcGHV-1 DNA was demonstrated in MAL-1 and HTRL. Furthermore, linear viral DNA was detected in MAL-1 cells as well as in TPA-treated HTRL (Fig. 5). It has been shown that the presence of circular viral genome is indicative of latent (non-productive) infection, whereas the occurrence of linear DNA indicates virus-producing cells [37]. Our results confirm published data that MAL-1 is productively infected with HVMA: previously, cell-free supernatant was used for a successful in vitro transformation of rabbit lymphocytes [19], but also for inoculation of rabbits followed by the development of tumours [21–23]. In contrast to MAL-1 cells, HTRL may be unable to produce viable virus. These cells might represent an example of an abortive infection mode after experimental transspecies infection as previously demonstrated for bovine herpesvirus 4 [68]. Restricted host range of MarcGHV-1 is further substantiated by the failure to experimentally infect human lymphocytes [19].

Most likely, DNA preparations of MAL-1 and HTRL do not contain proviral STLV-1 DNA. Three lines of evidence support this view. First, lack of proviral DNA was demonstrated by sensitive nested PCR (Fig. S1). Second, no relevant read pairs were found that map with STLV-1 and STLV-3 (partial) genomes available in GenBank. Third, rabbits which were inoculated with MAL-1 did not develop antibodies against closely related HTLV-1 [22]. Thus, we conclude that MarcGHV-1 rather than STLV-like retroviruses contributes to cellular transformation and to the development of tumours as is the case for EBV.

B- and T-lymphocyte-specific transcripts which were detected in MAL-1 and HTRL by Illumina sequencing (summarised in Table 2) indicate that both cell lines may represent mixed populations of B- and T-cells. This is in line with the known immunophenotype of MAL-1 which was previously demonstrated by cytometry [69] but was unclear for HTRL.

To better characterise the status of MarcGHV-1 in MAL-1 and HTRL, we have analysed the transcription and expression of EBV latency protein homologs. EBER-like reactivity was demonstrated in various nuclei of both cell lines (Table 3; Figs. 6c, d, 7). EBNA-1-like transcripts, known transactivators of EBV latency genes [70], were also detected in both cell lines. EBNA-1 of EBV inhibits proteasomal antigen degradation via its glycine–alanine repeats [46]. This mechanism seems to be specific for EBV and is lacking in simian LCVs [71]. The EBNA-1 gene of MarcGHV-1 encodes glycine–alanine–serine repeats.

The expression of at least eight EBV-like latency genes in MAL-1 cells is compatible with type III latency. However, in divergence to previous results [22], we could show a strong EBNA-2 expression by IHC, although the transcript was only weakly detectable. This discrepancy may be explained by the improved IHC technique used in this study which resulted in a perceivable increase in sensitivity. HTRL cells, however, lacked EBNA-2, -3A, and -3B transcripts but expressed low levels of EBNA-3C and high levels of LMP-1 and -2A-like transcripts. This finding prompted us to presume type II latency. In MAL-1, but not in HTRL, transcription and expression of ZEBRA (BZLF1) was detected. This protein triggers EBV reactivation [72]. Its expression was assessed as further confirmation of a productive MarcGHV-1 cycle in MAL-1. The absence of BZLF1 in HTRL is in accordance with the absence of linear DNA in native HTRL, as demonstrated by in situ lysis gel electrophoresis (Fig. 5); this confirms that HTRL is not productively infected by MarcGHV-1. In line with this, marked transcription of the early genes BSLF2/BMLF1 was only observed in MAL-1. Some slight discrepancies were seen between IHC and transcription analyses; however, it is well known that correlation between mRNA and their proteins sometimes differ

[73]. For example, in this study no LMP-1 expression was detectable by monoclonal antibodies which confirmed the results of a previous report [22]. However, LMP-1 transcript was verified in both cell lines.

MAL-1 cells have a stable, but massively rearranged, nearly-tetraploid macaque karyotype (Fig. S3). Furthermore, in vitro transformation of rabbit lymphocytes by MarcGHV-1 or even infection of rabbits resulted in the occurrence of chromosomal anomalies [19, 21]. The application of MarcGHV-1-specific FISH on MAL-1 and HTRL chromosomal preparations showed multiple inserts in both chromosomal preparations. Thus, chromosomal integration of MarcGHV-1 DNA was clearly demonstrated for both cell lines. The number and intensity of signals appeared to be higher in HTRL than in MAL-1 cells (see Fig. 8a, b). Furthermore, a mix of four probes representing different parts of the MarcGHV-1 genome showed co-localisation and clear evidence of chromosomal integration of complete virus genomes (Fig. 8c, d). Different MAL-1 metaphases showed that integration sites varied from cell to cell, but chromosomes #4, #6 and #16 were mostly affected by MarcGHV-1 and showed a similar localisation of DNA integration (Table S2 and Fig. S5). Chromosomal integration of MarcGHV-1 was further confirmed by Illumina sequencing data. Approximately 400 read pairs with joined overlaps between the host DNA and the viral genome were shown. Chromosomal integration of MarcGHV-1 DNA into various sites of the host genome resembles the situation observed in EBV-transformed cell lines [74–77]. Likewise, tumourigenesis in rabbits may arise from the expression of oncogenes and alterations of cellular gene expression [74].

Taken together, MarcGHV-1 is very closely related to EBV and LCVs of Old World monkeys. This gammaherpesvirus is able to infect rabbits and induce an abortive infection which may lead to lymphomas. Thus, MarcGHV-1 and its host cells might be a suitable candidate for an alternative EBV model. The MarcGHV-1-transformed cell lines MAL-1 and HTRL harbour B- and T-cell markers and carry integrated and episomal viral DNA. Both cell lines differ in their type of EBV-like latency, but only MAL-1 cells are able to produce viable virus.

Acknowledgements We thank the Institute of Clinical Molecular Biology (ICMB) in Kiel for providing Sanger sequencing facilities as supported in part by the DFG Cluster of Excellence "Inflammation at Interfaces" and "Future Ocean". We thank the technicians S. Greve, T. Henke, C. Noack, C. Botz von Drathen, M. Müller, I. Görlich and Dr. N. Kosyakova for excellent technical support. The authors would like to thank Dr. K. Korn (Virological Institute, University of Erlangen-Nuremberg, Erlangen, Germany) for providing a HTLV-1-positive DNA sample with known proviral load. Furthermore, authors are indebted to Dr. L. Harder (Institute of Tumour Genetics, Kiel, Germany) and R. Bunke, M.D. (Wietze, Germany) for critical reading of the manuscript and valuable comments.

Author contributions AK, RZ, HF, AS, AM, and PW planned the experiments; MS generated HTRL and performed initial characterisation; JR and MG performed sequencing with help of AK and RZ; these data were then analysed by JR, MG, AK, and RZ and used for phylogenetic comparisons; JR performed in situ lysis gel and Southern blot, and produced cloned MarcGHV-1 DNA for synthesis of FISH probes; chromosomal analysis and FISH were done by TL; AK, JR, and MG performed analysis of transcripts as well as STLV-1/HTLV-1 proviral DNA search; ISH and IHC were done by AK, GM and WK; AK, JR, TL and RZ wrote the paper; all authors approved the final version. Parts of this manuscript are subject of the MD thesis of JR.

Funding This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Compliance with ethical standards

Ethical statement Nothing to declare.

Conflict of interest The authors declare no financial conflict of interest and no non-financial conflict of interest.

References

1. Jha HC, Banerjee S, Robertson ES (2016) The role of gammaherpesviruses in cancer pathogenesis. Pathogens 5 (1). https://doi.org/10.3390/pathogens5010018
2. Thorley-Lawson DA, Gross A (2004) Persistence of the Epstein–Barr virus and the origins of associated lymphomas. N Engl J Med 350(13):1328–1337. https://doi.org/10.1056/NEJMra032015
3. Klein E, Kis LL, Klein G (2007) Epstein–Barr virus infection in humans: from harmless to life endangering virus-lymphocyte interactions. Oncogene 26(9):1297–1305. https://doi.org/10.1038/sj.onc.1210240
4. Hatton OL, Harris-Arnold A, Schaffert S, Krams SM, Martinez OM (2014) The interplay between Epstein–Barr virus and B lymphocytes: implications for infection, immunity, and disease. Immunol Res 58(2–3):268–276. https://doi.org/10.1007/s12026-014-8496-1
5. Young LS, Rickinson AB (2004) Epstein–Barr virus: 40 years on. Nat Rev Cancer 4(10):757–768. https://doi.org/10.1038/nrc1452
6. Stanfield BA, Luftig MA (2017) Recent advances in understanding Epstein–Barr virus. F1000Res 6:386. https://doi.org/10.12688/f1000research.10591.1
7. Shannon-Lowe C, Rickinson AB, Bell AI (2017) Epstein–Barr virus-associated lymphomas. Philos Trans R Soc Lond B Biol Sci 372 (1732). https://doi.org/10.1098/rstb.2016.0271
8. Krumbholz A, Sandhaus T, Gohlert A, Heim A, Zell R, Egerer R, Breuer M, Straube E, Wutzler P, Sauerbrei A (2010) Epstein–Barr virus-associated pneumonia and bronchiolitis obliterans syndrome in a lung transplant recipient. Med Microbiol Immunol 199(4):317–322. https://doi.org/10.1007/s00430-010-0165-y
9. Kamperschroer C, Gosink MM, Kumpf SW, O'Donnell LM, Tartaro KR (2016) The genomic sequence of lymphocryptovirus from cynomolgus macaque. Virology 488:28–36. https://doi.org/10.1016/j.virol.2015.10.025
10. Blossom D (2007) EBV and KSHV-related herpesviruses in non-human primates. In: Arvin A, Campadelli-Fiume G, Mocarski E et al (eds) Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press, Cambridge, pp 1093–1114

11. Ehlers B, Dural G, Yasmum N, Lembo T, de Thoisy B, Ryser-Degiorgis MP, Ulrich RG, McGeoch DJ (2008) Novel mammalian herpesviruses and lineages within the Gammaherpesvirinae: cospeciation and interspecies transfer. J Virol 82(7):3509–3516. https://doi.org/10.1128/JVI.02646-07
12. Pellett P, Davison A, Eberle R, Ehlers B, Hayward G, Lacoste V, Minson A, Nicholas J, Roizman B, Studdert M (2012) Order herpesvirales. Virus taxonomy: Classification and nomenclature of viruses: ninth report of the International Committee on Taxonomy of Viruses, King AMQ, Adams MJA, Carstens EB, Lefkowitz EJ, Academic Press, Waltham, pp 99–123
13. Rivailler P, Carville A, Kaur A, Rao P, Quink C, Kutok JL, Westmoreland S, Klumpp S, Simon M, Aster JC, Wang F (2004) Experimental rhesus lymphocryptovirus infection in immunosuppressed macaques: an animal model for Epstein–Barr virus pathogenesis in the immunosuppressed host. Blood 104(5):1482–1489. https://doi.org/10.1182/blood-2004-01-0342
14. Voevodin AF, Marx PA (2009) Lymphocryptoviruses. simian virology:323–346
15. Rivailler P, Jiang H, Cho YG, Quink C, Wang F (2002) Complete nucleotide sequence of the rhesus lymphocryptovirus: genetic validation for an Epstein–Barr virus animal model. J Virol 76(1):421–426
16. Wang F, Rivailler P, Rao P, Cho Y (2001) Simian homologues of Epstein–Barr virus. Philos Trans R Soc Lond B Biol Sci 356(1408):489–497. https://doi.org/10.1098/rstb.2000.0776
17. Cho Y, Ramer J, Rivailler P, Quink C, Garber RL, Beier DR, Wang F (2001) An Epstein–Barr-related herpesvirus from marmoset lymphomas. Proc Natl Acad Sci U S A 98(3):1224–1229. https://doi.org/10.1073/pnas.98.3.1224
18. Lapin BA, Timanovskaya VV, Yakovleva LA (1985) Herpesvirus HVMA: a new representative in the group of the EBV-like B-lymphotropic herpesviruses of primates. Haematol Blood Transfus 29:312–313
19. Meerbach A, Friedrichs C, Thust R, Wutzler P (2004) Transformation of rabbit lymphocytes by an Epstein–Barr virus-related herpesvirus from *Macaca arctoides*. Arch Virol 149(6):1083–1094. https://doi.org/10.1007/s00705-004-0293-z
20. Iakovleva LA, Timanovskaia VV, Indzhiia LV, Lapin BA, Voevodin AF (1987) Modelling of malignant lymphoma in rabbits using primate oncogenic viruses. Preliminary report. Biull Eksp Biol Med 103(3):336–338
21. Timanovskaia VV, Voevodin AF, Iakovleva LA, Markarian DS, Ivanov MT (1988) Malignant lymphoma in rabbits induced by the administration of herpes virus-containing material from brown macaques. Eksp Onkol 10(3):47–51
22. Wutzler P, Meerbach A, Farber I, Wolf H, Scheibner K (1995) Malignant lymphomas induced by an Epstein–Barr virus-related herpesvirus from Macaca arctoides—a rabbit model. Arch Virol 140(11):1979–1995
23. Yakovleva LA, Timanovskaya VV, Voevodin AF, Indzhiia LV, Lapin BA, Ivanov MT, Markaryan DS (1987) Modelling of malignant lymphoma in rabbits, using oncogenic viruses of non-human primates. Haematol Blood Transfus 31:445–447
24. Schatzl H, Tschikobava M, Rose D, Voevodin A, Nitschko H, Sieger E, Busch U, von der Helm K, Lapin B (1993) The Sukhumi primate monkey model for viral lymphomogenesis: high incidence of lymphomas with presence of STLV-I and EBV-like virus. Leukemia 7(Suppl 2):S86–S92
25. Pulvertaft RJV (1964) Cytology of Burkitt's tumor (African lymphoma). Lancet 1(7327):238–240
26. Karpova MB, Schoumans J, Ernberg I, Henter JI, Nordenskjold M, Fadeel B (2005) Raji revisited: cytogenetics of the original Burkitt's lymphoma cell line. Leukemia 19(1):159–161. https://doi.org/10.1038/sj.leu.2403534
27. Klein G, Giovanella B, Westman A, Stehlin JS, Mumford D (1975) An EBV-genome-negative cell line established from an American Burkitt lymphoma; receptor characteristics. EBV infectibility and permanent conversion into EBV-positive sublines by in vitro infection. Intervirology 5(6):319–334. https://doi.org/10.1159/000149930
28. Davies AH, Grand RJ, Evans FJ, Rickinson AB (1991) Induction of Epstein–Barr virus lytic cycle by tumor-promoting and non-tumor-promoting phorbol esters requires active protein kinase C. J Virol 65(12):6838–6844
29. Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357–359. https://doi.org/10.1038/nmeth.1923
30. d'Offay JM, Eberle R, Sucol Y, Schoelkopf L, White MA, Valentine BD, White GL, Lerche NW (2007) Transmission dynamics of simian T-lymphotropic virus type 1 (STLV1) in a baboon breeding colony: predominance of female-to-female transmission. Comp Med 57(1):105–114
31. Govert F, Krumbholz A, Witt K, Hopfner F, Feldkamp T, Korn K, Knoll A, Jansen O, Deuschl G, Fickenscher H (2015) HTLV-1 associated myelopathy after renal transplantation. J Clin Virol 72:102–105. https://doi.org/10.1016/j.jcv.2015.09.010
32. Dehee A, Cesaire R, Desire N, Lezin A, Bourdonne O, Bera O, Plumelle Y, Smadja D, Nicolas JC (2002) Quantitation of HTLV-I proviral load by a TaqMan real-time PCR assay. J Virol Methods 102(1–2):37–51
33. Edgar RC (2004) MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinf 5:113. https://doi.org/10.1186/1471-2105-5-113
34. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797. https://doi.org/10.1093/nar/gkh340
35. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol 30(12):2725–2729. https://doi.org/10.1093/molbev/mst197
36. Davison AJ (2010) Herpesvirus systematics. Vet Microbiol 143(1):52–69. https://doi.org/10.1016/j.vetmic.2010.02.014
37. Gardella T, Medveczky P, Sairenji T, Mulder C (1984) Detection of circular and linear herpesvirus DNA molecules in mammalian cells by gel electrophoresis. J Virol 50(1):248–254
38. Stuhlmann-Laeisz C, Szczepanowski M, Borchert A, Bruggemann M, Klapper W (2015) Epstein–Barr virus-negative diffuse large B-cell lymphoma hosts intra- and peritumoral B-cells with activated Epstein–Barr virus. Virchows Arch 466(1):85–92. https://doi.org/10.1007/s00428-014-1661-z
39. Stuhlmann-Laeisz C, Borchert A, Quintanilla-Martinez L, Hoeller S, Tzankov A, Oschlies I, Kreuz M, Trappe R, Klapper W (2016) In Europe expression of EBNA2 is associated with poor survival in EBV-positive diffuse large B-cell lymphoma of the elderly. Leuk Lymphoma 57(1):39–44. https://doi.org/10.3109/10428194.2015.1040014
40. Claussen U, Michel S, Muhlig P, Westermann M, Grummt UW, Kromeyer-Hauschild K, Liehr T (2002) Demystifying chromosome preparation and the implications for the concept of chromosome condensation during mitosis. Cytogenet Genome Res 98(2–3):136–146. https://doi.org/10.1159/000069817
41. Liehr T, Weise A, Heller A, Starke H, Mrasek K, Kuechler A, Weier HU, Claussen U (2002) Multicolor chromosome banding (MCB) with YAC/BAC-based probes and region-specific microdissection DNA libraries. Cytogenet Genome Res 97(1–2):43–50. https://doi.org/10.1159/000064043
42. Liehr T, Weise A, Hamid AB, Fan X, Klein E, Aust N, Othman MA, Mrasek K, Kosyakova N (2013) Multicolor FISH methods in current clinical diagnostics. Expert Rev Mol Diagn 13(3):251–255. https://doi.org/10.1586/erm.12.146

43. Weise A, Gross M, Mrasek K, Mkrtchyan H, Horsthemke B, Jonsrud C, Von Eggeling F, Hinreiner S, Witthuhn V, Claussen U, Liehr T (2008) Parental-origin-determination fluorescence in situ hybridization distinguishes homologous human chromosomes on a single-cell level. Int J Mol Med 21(2):189–200
44. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP,Evers DJ, Barnes CL, Bignell HR, Boutell JM, Bryant J, Carter RJ, Keira Cheetham R,Cox AJ, Ellis DJ, Flatbush MR, Gormley NA, Humphray SJ, Irving LJ, Karbelashvili MS,Kirk SM, Li H, Liu X, Maisinger KS, Murray LJ, Obradovic B, Ost T, Parkinson ML, Pratt MR, Rasolonjatovo IM, Reed MT, Rigatti R, Rodighiero C, Ross MT, Sabot A, Sankar SV, Scally A, Schroth GP, Smith ME, Smith VP, Spiridou A, Torrance PE, Tzonev SS, Vermaas EH, Walter K, Wu X, Zhang L, Alam MD, Anastasi C, Aniebo IC, Bailey DM, Bancarz IR, Banerjee S, Barbour SG, Baybayan PA, Benoit VA, Benson KF, Bevis C, Black PJ, Boodhun A, Brennan JS, Bridgham JA, Brown RC, Brown AA, Buermann DH, Bundu AA, Burrows JC,Carter NP, Castillo N, Chiara ECM, Chang S, Neil Cooley R, Crake NR, Dada OO, Diakoumakos KD, Dominguez-Fernandez B, Earnshaw DJ, Egbujor UC, Elmore DW, Etchin SS, Ewan MR,Fedurco M, Fraser LJ, Fuentes Fajardo KV, Scott Furey W, George D, Gietzen KJ, Goddard CP, Golda GS, Granieri PA, Green DE, Gustafson DL, Hansen NF, Harnish K, Haudenschild CD, Heyer NI, Hims MM, Ho JT, Horgan AM, Hoschler K, Hurwitz S, Ivanov DV, Johnson MQ, James T, Huw Jones TA, Kang GD, Kerelska TH, Kersey AD, Khrebtukova I, Kindwall AP, Kingsbury Z, Kokko-Gonzales PI, Kumar A, Laurent MA, Lawley CT, Lee SE, Lee X,Liao AK, Loch JA, Lok M, Luo S, Mammen RM, Martin JW, McCauley PG, McNitt P, Mehta P, Moon KW, Mullens JW, Newington T, Ning Z, Ling Ng B, Novo SM, O'Neill MJ, Osborne MA, Osnowski A, Ostadan O, Paraschos LL, Pickering L, Pike AC, Pike AC, Chris Pinkard D, Pliskin DP, Podhasky J, Quijano VJ, Raczy C, Rae VH, Rawlings SR, Chiva Rodriguez A, Roe PM, Rogers J, Rogert Bacigalupo MC, Romanov N, Romieu A, Roth RK, Rourke NJ,Ruediger ST, Rusman E, Sanches-Kuiper RM, Schenker MR, Seoane JM, Shaw RJ, Shiver MK, Short SW, Sizto NL, Sluis JP, Smith MA, Ernest Sohna Sohna J, Spence EJ, Stevens K, Sutton N, Szajkowski L, Tregidgo CL, Turcatti G, Vandevondele S, Verhovsky Y, Virk SM, Wakelin S, Walcott GC, Wang J, Worsley GJ, Yan J, Yau L, Zuerlein M, Rogers J,Mullikin JC, Hurles ME, McCooke NJ, West JS, Oaks FL, Lundberg PL, Klenerman D, Durbin R, Smith AJ (2008) Accurate whole human genome sequencing using reversible terminator chemistry. Nature 456(7218):53–59. https://doi.org/10.1038/nature07517
45. Ning S (2011) Innate immune modulation in EBV infection. Herpesviridae 2(1):1. https://doi.org/10.1186/2042-4280-2-1
46. Levitskaya J, Coram M, Levitsky V, Imreh S, Steigerwald-Mullen PM, Klein G, Kurilla MG, Masucci MG (1995) Inhibition of antigen processing by the internal repeat region of the Epstein–Barr virus nuclear antigen-1. Nature 375(6533):685–688. https://doi.org/10.1038/375685a0
47. Roizman B, Baines J (1991) The diversity and unity of herpesviridae. Comp Immunol Microbiol Infect Dis 14(2):63–79
48. Le SQ, Gascuel O (2008) An improved general amino acid replacement matrix. Mol Biol Evol 25(7):1307–1320. https://doi.org/10.1093/molbev/msn067
49. McKillen J, Hogg K, Lagan P, Ball C, Doherty S, Reid N, Collins L, Dick JTA (2017) Detection of a novel gammaherpesvirus (genus Rhadinovirus) in wild muntjac deer in Northern Ireland. Arch Virol 162(6):1737–1740. https://doi.org/10.1007/s00705-017-3254-z
50. Nei M, Kumar S (2000) Molecular evolution and phylogenetics. Oxford University Press, New York
51. Jackson SA, DeLuca NA (2003) Relationship of herpes simplex virus genome configuration to productive and persistent infections. Proc Natl Acad Sci USA 100(13):7871–7876. https://doi.org/10.1073/pnas.1230643100
52. Sandri-Goldin RM (2003) Replication of the herpes simplex virus genome: does it really go around in circles? Proc Natl Acad Sci USA 100(13):7428–7429. https://doi.org/10.1073/pnas.1432875100
53. Chau CM, Zhang XY, McMahon SB, Lieberman PM (2006) Regulation of Epstein–Barr virus latency type by the chromatin boundary factor CTCF. J Virol 80(12):5723–5732. https://doi.org/10.1128/JVI.00025-06
54. Hurley EA, Agger S, McNeil JA, Lawrence JB, Calendar A, Lenoir G, Thorley-Lawson DA (1991) When Epstein–Barr virus persistently infects B-cell lines, it frequently integrates. J Virol 65(3):1245–1254
55. Watanabe T, Seiki M, Tsujimoto H, Miyoshi I, Hayami M, Yoshida M (1985) Sequence homology of the simian retrovirus genome with human T-cell leukemia virus type I. Virology 144(1):59–65
56. Bushman F, Lewinski M, Ciuffi A, Barr S, Leipzig J, Hannenhalli S, Hoffmann C (2005) Genome-wide analysis of retroviral DNA integration. Nat Rev Microbiol 3(11):848–858. https://doi.org/10.1038/nrmicro1263
57. Pezzutto A, Ulrichs T, Burmester G-R (2007) Taschenatlas der Immunologie, vol 2. Georg Thieme Verlag, Stuttgart
58. Ok CY, Li L, Young KH (2015) EBV-driven B-cell lymphoproliferative disorders: from biology, classification and differential diagnosis to clinical management. Exp Mol Med 47:e132. https://doi.org/10.1038/emm.2014.82
59. Wienberg J, Stanyon R, Jauch A, Cremer T (1992) Homologies in human and Macaca fuscata chromosomes revealed by in situ suppression hybridization with human chromosome specific DNA libraries. Chromosoma 101(5–6):265–270
60. Fan X, Sangpakdee W, Tanomtong A, Chaveerach A, Pinthong K, Pornnarong S, Supiwong W, Trifonov V, Hovhannisyan G, Aroutouinian R (2014) Molecular cytogenetic analysis of Thai southern pig-tailed macaque (*Macaca nemestrina*) by multicolor banding. Proceedings of Yerevan State University 2014:46–50
61. Fan X, Sangpakdee W, Tanomtong A, Chaveerach A, Pinthong K, Pornnarong S, Supiwong W, Trifonov V, Hovhannisyan G, Loth K (2014) Comprehensive molecular cytogenetic analysis of Barbary macaque (*Macaca sylvanus*). J Armenia 66(1):98–102
62. O'Brien SJ, Menninger JC, Nash WG (2006) Atlas of mammalian chromosomes. Wiley, Hoboken
63. Ono K, Satoh M, Yoshida T, Ozawa Y, Kohara A, Takeuchi M, Mizusawa H, Sawada H (2007) Species identification of animal cells by nested PCR targeted to mitochondrial DNA. In Vitro Cell Dev Biol Anim 43(5–6):168–175. https://doi.org/10.1007/s11626-007-9033-5
64. Ahmed EH, Baiocchi RA (2016) Murine models of Epstein–Barr virus-associated lymphomagenesis. ILAR J 57(1):55–62. https://doi.org/10.1093/ilar/ilv074
65. Ito R, Takahashi T, Katano I, Ito M (2012) Current advances in humanized mouse models. Cell Mol Immunol 9(3):208–214. https://doi.org/10.1038/cmi.2012.2
66. Gujer C, Chatterjee B, Landtwing V, Raykova A, McHugh D, Munz C (2015) Animal models of Epstein Barr virus infection. Curr Opin Virol 13:6–10. https://doi.org/10.1016/j.coviro.2015.03.014
67. Kakubava VV, Agrba VZ, Timanovskaia VV, Brzhikhachek B (1990) Analysis of B-lymphotropic oncogenic herpesvirus of *Macaca arctoides*. Eksp Onkol 12(6):44–46
68. Gillet L, Minner F, Detry B, Farnir F, Willems L, Lambot M, Thiry E, Pastoret PP, Schynts F, Vanderplasschen A (2004) Investigation of the susceptibility of human cell lines to bovine

herpesvirus 4 infection: demonstration that human cells can support a nonpermissive persistent infection which protects them against tumor necrosis factor alpha-induced apoptosis. J Virol 78(5):2336–2347
69. Agrba VZ, Lapin BA, Medvedeva NM, Yakovleva LA (2004) Immunophenotypical characteristics of permanent cultures of lymphoid cells from Papio hamadryas and *Macaca arctoides*. Bull Exp Biol Med 137(2):190–194
70. Kang MS, Kieff E (2015) Epstein–Barr virus latent genes. Exp Mol Med 47:e131. https://doi.org/10.1038/emm.2014.84
71. Blake NW, Moghaddam A, Rao P, Kaur A, Glickman R, Cho YG, Marchini A, Haigh T, Johnson RP, Rickinson AB, Wang F (1999) Inhibition of antigen presentation by the glycine/alanine repeat domain is not conserved in simian homologues of Epstein–Barr virus nuclear antigen 1. J Virol 73(9):7381–7389
72. Murata T (2014) Regulation of Epstein–Barr virus reactivation from latency. Microbiol Immunol 58(6):307–317. https://doi.org/10.1111/1348-0421.12155
73. Haider S, Pal R (2013) Integrated analysis of transcriptomic and proteomic data. Curr Genomics 14(2):91–110. https://doi.org/10.2174/1389202911314020003
74. Morissette G, Flamand L (2010) Herpesviruses and chromosomal integration. J Virol 84(23):12100–12109. https://doi.org/10.1128/JVI.01169-10
75. Lestou VS, De Braekeleer M, Strehl S, Ott G, Gadner H, Ambros PF (1993) Non-random integration of Epstein–Barr virus in lymphoblastoid cell lines. Genes Chromosomes Cancer 8(1):38–48
76. Caporossi D, Vernole P, Porfirio B, Tedeschi B, Frezza D, Nicoletti B, Calef E (1988) Specific sites for EBV association in the Namalwa Burkitt lymphoma cell line and in a lymphoblastoid line transformed in vitro with EBV. Cytogenet Cell Genet 48(4):220–223. https://doi.org/10.1159/000132632
77. Gao J, Luo X, Tang K, Li X, Li G (2006) Epstein–Barr virus integrates frequently into chromosome 4q, 2q, 1q and 7q of Burkitt's lymphoma cell line (Raji). J Virol Methods 136(1–2):193–199. https://doi.org/10.1016/j.jviromet.2006.05.013
78. Davison AJ (2007) Comparative analysis of the genomes. In: Arvin A, Campadelli-Fiume G, Mocarski E et al (eds) Human herpesviruses—biology, therapy and immunoprophylaxis. Cambridge University Press, Cambridge, pp 10–26
79. Young LS, Arrand JR, Murray PG (2007) EBV gene expression and regulation. In: Arvin A, Campadelli-Fiume G, Mocarski E et al (eds) Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge University Press, Cambridge, pp 461–489
80. Aubry V, Mure F, Mariame B, Deschamps T, Wyrwicz LS, Manet E, Gruffat H (2014) Epstein–Barr virus late gene transcription depends on the assembly of a virus-specific preinitiation complex. J Virol 88(21):12825–12838. https://doi.org/10.1128/JVI.02139-14
81. Gruffat H, Kadjouf F, Mariame B, Manet E (2012) The Epstein–Barr virus BcRF1 gene product is a TBP-like protein with an essential role in late gene expression. J Virol 86(11):6023–6032. https://doi.org/10.1128/JVI.00159-12
82. Fossum E, Friedel CC, Rajagopala SV, Titz B, Baiker A, Schmidt T, Kraus T, Stellberger T, Rutenberg C, Suthram S, Bandyopadhyay S, Rose D, von Brunn A, Uhlmann M, Zeretzke C, Dong YA, Boulet H, Koegl M, Bailer SM, Koszinowski U, Ideker T, Uetz P, Zimmer R, Haas J (2009) Evolutionarily conserved herpesviral protein interaction networks. PLoS Pathog 5(9):e1000570. https://doi.org/10.1371/journal.ppat.1000570
83. Chiu SH, Wu MC, Wu CC, Chen YC, Lin SF, Hsu JT, Yang CS, Tsai CH, Takada K, Chen MR, Chen JY (2014) Epstein–Barr virus BALF3 has nuclease activity and mediates mature virion production during the lytic cycle. J Virol 88(9):4962–4975. https://doi.org/10.1128/JVI.00063-14
84. Thompson MP, Kurzrock R (2004) Epstein–Barr virus and cancer. Clin Cancer Res 10(3):803–821