Complete Genomic Sequence of an Epstein-Barr Virus-Related Herpesvirus Naturally Infecting a New World Primate: a Defining Point in the Evolution of Oncogenic Lymphocryptoviruses

Pierre Rivaller,* Young-gyu Cho, and Fred Wang

Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115

Received 18 June 2002/Accepted 30 August 2002

Callitrichine herpesvirus 3 (CalHV-3) was isolated from a B-cell lymphoma arising spontaneously in the New World primate *Callithrix jacchus*, the common marmoset. Partial genomic sequence analysis definitively identified CalHV-3 as a member of the Epstein-Barr virus (EBV)-related lymphocryptovirus (LCV) genus and extended the known host range of LCVs beyond humans and Old World nonhuman primates. We have now completed the first genomic sequence of an LCV infecting a New World primate by describing the unique short region, the major internal repeat, and a portion of the unique long region. This portion of the genome contains the putative latent origin of replication and 13 additional open reading frames (ORFs), 5 of which show no homology to any viral or cell genes. One of the novel genes, C5, is a positional homologue for the transformation-essential EBV gene EBNA-2. The marmoset LCV genome is also notable for the absence of viral interleukin-10 and small nonpolyadenylated RNA homologues. Marmoset LCV transcripts encoding putative latent infection nuclear proteins have a common leader sequence that is spliced from the major internal repeat in a manner similar to that of the EBV EBNA-LP, suggesting strong conservation of a common promoter and splicing of these latent infection mRNAs. An EBV LMP2A-like spliced transcript crossing the terminal repeats encodes a unique ORF, C7, with multiple transmembrane domains and tyrosine kinase phosphorylation sites functionally reminiscent of EBV LMP2A. However, the carboxy-terminal location of the candidate phosphotyrosine residues is more reminiscent of the Kaposi's sarcoma-associated herpesvirus K15 gene and provides potential evidence of an evolutionary transition from rhadinoviruses to lymphocryptoviruses. The unusual gene repertoire of the marmoset LCV differentiates ancestral viral genes likely present in an LCV progenitor from viral genes acquired later as primates and LCV coevolved, providing a defining point in the evolution of oncogenic LCVs.

Epstein-Barr virus (EBV) is the prototypical gamma-1 herpesvirus (*Lymphocryptovirus* genus). EBV infects most adult humans, persists asymptomatically for life, and is associated with a variety of human diseases, including infectious mononucleosis, B-cell malignancies, and epithelial cell malignancies (for a review, see reference 47). EBV encodes more than 85 genes, and understanding how this repertoire of viral genes contributes to successful EBV infection and associated human diseases is a fundamental component of EBV research. In vitro studies have been useful for characterizing many EBV genes, including latent infection genes associated with B-cell transformation and lytic infection genes involved in viral replication and assembly (for a review, see reference 28). However, for 30 to 40% of EBV genes, very little is known.

Studies of EBV-related herpesviruses can provide new insights by comparison to EBV and through new animal model systems to study EBV pathogenesis. Lymphocryptoviruses (LCVs) naturally infecting Old World primates are known to be biologically similar to EBV (for a review, see reference 66). LCV infection is ubiquitous in adult Old World nonhuman primates, and these animals harbor persistent LCV infection in their peripheral blood. Old World LCVs can immortalize B cells in tissue culture, and LCV infection is associated with

tumorigenesis in vivo. We showed that Old World rhesus macaques could be used as an animal model for EBV infection, as experimental infection of naive rhesus macaques with rhesus LCV reproduced acute and persistent infection similar to EBV infection in humans (35). The complete genome sequence of rhesus LCV was derived as a prototype for Old World LCVs, and the identical gene repertoire and high degree of amino acid homology (75%) provided genetic validation for the similarities between EBV and rhesus LCV infection (48).

We recently identified the first EBV-related herpesvirus naturally infecting a New World primate (6). This transforming virus, isolated from a spontaneous B-cell lymphoma (43) arising in a common marmoset (*Callithrix jacchus*), was formally named Callitrichine herpesvirus 3 as the third herpesvirus described naturally infecting a member of the *Callitrichidae* family of New World primates. Cloning and sequencing from 105 kb of viral DNA revealed colinear genome organization with other gammaherpesviruses and 60 open reading frame (ORFs) that were more closely related to EBV and other LCVs than any other herpesviruses.

Further studies demonstrated that persistent Callitrichine herpesvirus 3 infection was prevalent in two different marmoset colonies and that a closely related virus could be detected in squirrel monkeys (*Saimiri sciureus*), a primate from the other major family of New World primates, *Cebidae* (6). These studies clearly established Callitrichine herpesvirus 3 as a member of the LCV genus, herein referred to as the marmoset LCV,

* Corresponding author. Mailing address: Channing Laboratory, 181 Longwood Ave., Boston, MA 02115. Phone: (617) 525-4226. Fax: (617) 525-4251. E-mail: privaller@rics.bwh.harvard.edu.

and demonstrated that the natural LCV host range includes humans and both Old and New World nonhuman primates.

The current study now completes the genome sequence for a prototypic New World LCV and provides an opportunity for a comparative analysis with the complete genome sequences of an Old World LCV (rhesus LCV), and the human LCV (EBV) to better understand the evolution of this oncogenic herpesvirus genus. All viruses in the LCV genus are capable of immortalizing B cells *in vitro* and are associated with persistent infection and B-cell lymphomagenesis in the natural host, indicating a strong biological selection for these properties throughout the evolution of this virus genus. These viruses most likely coevolved with their natural hosts, so that the Old World LCV evolved approximately 25 million years before EBV, whereas the New World LCV provides a glimpse of LCV evolution approximately 35 million years before the appearance of human LCV (20, 59). Thus, it is not surprising to find that the rhesus LCV genome is more closely related to EBV and that the marmoset LCV genome is more divergent from EBV.

The similarities and dramatic differences between the marmoset LCV and the rhesus LCV and EBV can distinguish between the genetic strategies and biological properties that may be fundamental to this genus versus features that have been acquired later in evolution as the host species evolved. In this way, completion of the marmoset LCV genome, definition of the viral gene repertoire, and initial descriptions of the complex transcription patterns can provide a defining point in the evolution of this oncogenic, EBV-related herpesvirus genus.

MATERIALS AND METHODS

Viral DNA cloning and sequencing. Genomic DNA from the marmoset LCV-infected cell line CJ0149 (6) was partially digested with *Sau3AI* and cloned into the *Bam*HI site of the Supercos I (Stratagene) vector. Viral DNA cosmids A10, D6, and D4 have been described previously (6). The overlapping cosmid E4 and plasmid H5 were isolated from the same library to extend the genomic sequence. The nucleotide coordinates for each viral DNA clone are as follows: E4, from the terminal repeat (TR) to nucleotide (nt) 119700; H5, nt 119891 to 105765; D4, nt 106440 to 70700; D6, nt 75819 to 36591; and A10, nt 40053 to the TR. Cosmid clones were sequenced by shotgun cloning of random cosmid DNA fragments (2 to 4 kb) derived by sonication. Overlapping nucleotide sequences were aligned with Sequencher software (GeneCode, version 4.05).

Nucleotide and amino acid sequence analysis. The marmoset LCV gene repertoire was analyzed with MapDraw (DNASTar). ORFs encoding more than 150 amino acids or homologous to known genes were characterized further. Marmoset LCV ORFs were identified by Blast analysis, and amino acid homology of marmoset LCV ORFs to the EBV, Kaposi's sarcoma-associated herpesvirus, and herpesvirus saimiri counterparts was determined with MegAlign (DNASTar). Spliced genes were identified with Gene Finder software (<http://dot.imgen.bcm.tmc.edu:9331/gene-finder/gf.html>). Transmembrane domains and nuclear localization signals were predicted with the servers at <http://sosui.proteome.bio.tuat.ac.jp/sosui/menu0.html> and <http://psort.ims.u-tokyo.ac.jp/form2.html>. Src homology-2 (SH2) domains were identified with the Scansite website (<http://scansite.mit.edu>), and the serine residues likely to be phosphorylated were mapped with the NetPhos 2.0 prediction server (<http://www.cbs.dtu.dk/services/NetPhos>).

RT-PCR analysis. Total RNA (2 µg) isolated from the marmoset LCV-infected cell line CJ0149 (6) was treated with DNase I for 15 min at room temperature, denatured for 10 min at 65°C and hybridized with random hexamers (25 pmol) for 10 min at 65°C. cDNAs were generated with Superscript II reverse transcriptase (RT; Invitrogen) for 1 h at 37°C. PCR primers used for C7 and C0 to C5/C3/Orf39 amplifications were as follows: C7E2F (5'-TCCTGGC TTATTTTTGTGCC-3') paired with C7E4R (5'-GTGGCAGCCAAAGGAAA ATA-3') and C0E3F (5'-CGTGTACGCACAAGGAGAA-3') paired with C5R (5'-AATCCGCCACACAGTCTAC-3'), C3R (5'-AGATGTGCTTGAACGC TCTG-3'), or 39R (5'-CTTTTCTGTATCCCGGCCT-3').

Rapid amplification of cDNA ends (RACE) analyses were performed with the

Smart Race cDNA amplification kit (Clontech). Reverse primers C7E4R and C7E2R (5'-AACACCAAAAACAAACGGC-3') were used for 5'RACE of the C7 transcript. Reverse primers C5R and C0E3R (5'-ATTCTCCTTGTGCGTA CACG-3') were used for 5'RACE on the C5 transcript. Forward primers C7E2F and C7E4F (5'-TCGGTATCTGCATTTTGGGA-3') were used for 3'RACE on the C7 transcript. PCR products were run on a 2% agarose gel and cloned into the TA vector (Invitrogen), and at least five clones from each PCR assay were sequenced.

Northern blot analysis. Total RNA (1 to 20 µg) isolated from marmoset LCV-infected CJ0149 cells (6), rhesus LCV-infected LCL8664 cells (44), or EBV-negative BJAB cells was separated by electrophoresis in a 1% formaldehyde-agarose gel. After transfer to a nylon membrane, RNA was hybridized with cosmid viral DNA that had been sheared and labeled with ³²P after priming with random hexamers. Recombinant human small nonpolyadenylated EBV-encoded RNAs (EBERs) were detected with the rhesus LCV cosmid CD1 (49), and RNA from marmoset LCV-infected cells was hybridized with marmoset LCV cosmid DNAs labeled in parallel. Blots were hybridized at 68°C in QuikHyb solution (Stratagene) and washed twice for 15 min at 25°C in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% (wt/vol) sodium dodecyl sulfate (SDS) and then for 30 min at 60°C in 0.1× SSC–0.1% (wt/vol) SDS.

Nucleotide sequence accession number. The complete nucleotide sequence of the marmoset LCV has been deposited in GenBank as an update of the previously reported sequence AF319782 (6).

RESULTS

Completion of marmoset LCV sequence. The remaining 50 kb of marmoset LCV sequence was derived from overlapping cosmid and plasmid viral DNA clones. This sequence corresponds to approximately one third of the EBV genome (nucleotides 1 to 68000 of the B95-8 genome) containing the short unique region (U_S), the major internal repeat (IR1), and a small portion of the long unique region (U_L) (Fig. 1A).

Most of the marmoset LCV U_L, including 60 ORFs, was described previously (6). The current data complete the marmoset LCV U_L by identifying nine additional ORFs (Fig. 1B). Seven of these ORFs (ORFs 58 to 64) had 20 to 52% amino acid homology to EBV lytic infection ORFs at a similar position in the EBV U_L (Table 1). Two additional marmoset LCV ORFs in the U_L had no sequence homology to cell or viral genes present in the GenBank database. These are identified as C4 and C5 according to the previous nomenclature used to identify unique marmoset LCV genes (6). C4 is a positional homologue of BHLF1, an EBV lytic infection gene of unknown function, and C5 is a positional homologue of the essential EBV transforming gene EBNA-2.

The marmoset LCV U_S is approximately 10 kb long and predicted to encode three ORFs (Fig. 1B). ORF65 is homologous to the EBV lytic gene BNRF1, with 41% amino acid homology. The other two ORFs had no sequence homology to cell or viral genes and are identified as C6 and C7.

Marmoset LCV U_L: C5, a positional homologue for the essential EBV transforming gene EBNA-2. The C5 product is predicted to be a 420-amino-acid protein with a nuclear localization signal, indicating that C5 is likely to reside in the cell nucleus, similar to EBNA-2 (Fig. 2). Despite no significant primary amino acid homology to EBNA-2, C5 had a cluster of acidic residues at the carboxy terminus, suggesting functional similarities to the EBNA-2 transcriptional transactivating domain (9). A striking difference in C5 was the lack of a polyproline repeat region that was present in both the type 1 and type 2 EBNA-2s from EBV and rhesus LCV (7, 40, 69). Thus, the polyproline repeat appears relatively late in LCV evolution, i.e., after the evolution of New World LCV.

A EBV

B Marmoset LCV

FIG. 1. Partial map of EBV and marmoset LCV genomes, showing the U_S, IR1, and a portion of the U_L region. (A) Map of EBV U_S, IR1, and a portion of U_L. The map is drawn to scale, and the nucleotide coordinates as numbered in GenBank are shown at the top. U_S and U_L regions are delimited by arrows, and repeat regions are boxed. The lytic origin of replication (OriP) is shown, as well as the latent origin of replication (OriLyt) with the family of repeats (FR) and dyad symmetry element (DS). Each ORF is represented as a black arrow. The orientation of the ORFs is shown by the direction of the arrow. EBERs are represented by short lines. EBNA-LP and EBNA-2 are also indicated. (B) Map of marmoset LCV U_S, IR1, and a portion of U_L. Each ORF is represented as an open arrow.

The polyproline repeat is much shorter in the rhesus and baboon LCV than in EBV EBNA-2 (27, 20, and 40 proline residues, respectively) (40), suggesting that the size of the polyproline repeat may be increasing in parallel with human evolution. The majority of the EBV EBNA-2 polyproline repeat can be deleted without adverse effects on EBNA-2 transcriptional transactivating and B-cell-immortalizing properties (67), indicating that the polyproline repeat is not critical for EBNA-2's growth-transforming function. The role of the EBNA-2 polyproline repeat may be more subtle, providing some biological advantage for virus infection *in vivo* that is not apparent in tissue culture systems.

Marmoset LCV U_S: C7, a highly spliced transcript across the terminal repeats. Computer-assisted analysis suggested the presence of a multiply spliced transcript in the marmoset LCV U_S between the TR and the BNRF1 homologue ORF65. Initial RT-PCR analysis confirmed the expression of a transcript with three exons coding for multiple hydrophobic transmembrane domains. Six additional exons were identified by 3' RACE, and one additional exon was identified by 5' RACE. The 5'-most exon was located on the other side of the TR, similar to the EBV LMP2A and LMP2B transcripts (Fig. 1B). However, in contrast to the EBV LMP2A and -2B first exons, located downstream and upstream of LMP1, respectively, this exon was located in a C1 intron.

Among six independent 5'RACE clones sequenced, the three longest clones started at the same 5' nucleotide, and there was a potential initiator methionine codon at nucleotide

94 of these RT-PCR products (Fig. 3A). However, this site did not fit a consensus Kozak sequence for translational initiation, and no obvious promoter element was present immediately upstream. The amino acid sequence at the 5' end stayed in frame with the longest ORF, and additional 5' sequence or exons cannot be ruled out.

The longest ORF encoded a 413-amino-acid protein with no sequence homology to known cell or viral genes and was therefore identified as C7 (Fig. 3A). Like LMP2, it had 12 transmembrane domains, but the putative C7 peptide sequence differed from LMP2 by a shorter amino terminus and a longer carboxy terminus. In addition, three of the five tyrosine residues in the C7 carboxy-terminal domain were predicted to be phosphorylated as Src homology-2 (SH2) domains (Fig. 3A and 3B). This is reminiscent of the LMP2A amino terminus, which contains eight tyrosine residues with at least four SH2 domains and one immunoreceptor tyrosine-based activation motif (33). EBV LMP2A can block signaling through the B-cell receptor by recruiting "Src family" tyrosine kinases (5, 12, 14, 15). The presence of SH2 domains within the C7 cytoplasmic tail suggests that C7 might have a function similar to that of EBV LMP2A. However, the overall structure of C7 with SH2 domains in a longer carboxy terminus is more similar to the K15 membrane protein encoded by a highly spliced gene that is near the termini of the human rhadinovirus Kaposi's sarcoma-associated herpesvirus (Fig. 3B) (8, 17). Thus, C7 may represent lingering evidence of an evolutionary transition from rhadinoviruses to lymphocryptoviruses.

ORF15	L	35887	34673	404	BTRF1	425	57.0	ORF23	24.8	ORF23	19.8	Major capsid protein (p135) Capsid protein gp150 Packaging protein
ORF16	L	38183	35802	793	BcRF1	575	44.2	ORF24	25.8	ORF24	25.8	
ORF17	R	38182	42321	1,380	BcLF1	1,381	75.5	ORF25	55.7	ORF25	55.5	
ORF18	R	42337	43242	302	BDLF1	301	70.1	ORF26	50.0	ORF26	49.3	
ORF19	R	43244	44419	392	BDLF2	420	26.3	ORF27	13.1	ORF27	16.1	
ORF20	L	46106	45054	351	BDLF3	234	74.6	ORF28	57.8	ORF28	53.8	Packaging protein
ORF21	R	46603	47238	212	BDRF1	379	51.9	ORF29b	28.3	ORF29b	29.8	
ORF22	R	47208	48734	509	BDLF4	225	30.6	ORF30	15.0	ORF30	20.9	
ORF23	R	48685	49695	337	BGLF1	507	57.4	ORF31	24.5	ORF31	30.9	
ORF24	L	50575	49874	234	BGLF2	336	63.7	ORF32	45.3	ORF32	47.0	
ORF25	R	50859	51938	360	BGRF1	311	48.2	ORF29a	30.0	ORF29a	31.6	Kinase Alkaline exonuclease Tegument protein Glycoprotein M Helicase-primase complex Helicase-primase complex
ORF26	R	51925	52371	149	BGLF3.5	332	35.6	ORF33	26.2	ORF33	21.5	
ORF27	R	52274	53620	449	BGLF4	455	59.9	ORF34	24.1	ORF34	24.4	
ORF28	R	53749	55161	471	BGLF5	470	69.6	ORF35	36.9	ORF35	40.6	
ORF29	R	55116	55343	78	BBLF1	75	38.7	ORF36	16.4	ORF36	22.7	
ORF30	L	59562	58408	385	BBRF3	405	55.6	ORF37	40.5	ORF37	43.2	Capsid protein Helicase-primase complex
ORF31	R	59695	61290	532	BBRF2	522	20.3	ORF38	13.6	ORF38	13.1	
ORF32	R	61536	62138	201	BBLF3	201	21.9	ORF39	14.4	ORF39	16.1	
ORF33	L	62977	62135	281	BBRF2	278	55.8	ORF40	27.7	ORF40	30.6	
ORF34	L	64697	62901	599	BBRF1	613	70.6	ORF41	46.6	ORF41	49.9	
ORF35	R	64627	67113	827	BBLF4	809	65.3	ORF42	47.8	ORF42	48.4	Uracil DNA glycosylase Glycoprotein L EBNA-1 homologue
ORF36	L	67704	67132	191	BKRF4	217	27.7	ORF43	15.2	ORF43	17.8	
ORF37	L	68482	67712	257	BKRF3	255	74.9	ORF44	46.7	ORF44	50.0	
ORF38	L	68880	68449	124	BKRF2	137	50.4	ORF45	18.1	ORF45	22.7	
ORF39	L	70133	69150	328	BKRF1	641	36.0	ORF46	14.2	ORF46	12.2	
ORF40	L	71603	70182	474	BRRF2	537	22.4	ORF47	16.2	ORF47	14.5	Transactivator Transactivator Transactivator
ORF41	L	72691	71780	304	BRRF1	310	46.7	ORF48	13.8	ORF48	15.5	
ORF42	R	72699	74525	609	BRLF1	605	39.0	ORF49		ORF49		
ORF43	R	74709	75224	252	BZLF1	245	29.0	ORF50		ORF50		
ORF44	R	75358	75465	36	BZLF2	223	33.6	ORF51		ORF51		
C3	L	75702	75836	44								Nuclear protein Nuclear protein Nuclear protein gp350 dUTPase DNA replication protein Transactivator vIRF1
ORF45	R	75840	76565	242								
ORF46	L	79013	76593	825	EBNA-3C	1,069						
ORF47	L	79162	79106	18	EBNA-3B	938						
ORF48	R	79328	81427	700	EBNA-3A	925	22.4					
ORF49	L	81945	81469	159	gp350	886						Membrane protein
ORF50	R	82327	83220	298	BLF2	148						
ORF51	R	84138	83479	220	BLRF2	162	40.3	ORF52	22.1	ORF52	23.5	
		84182	86686	835	BLRF1	102	44.1	ORF53	34.6	ORF53	27.8	
		86943	88178	412	BBLF3	102	54.3	ORF54	26.8	ORF54	23.7	
ORF52	L	89773	88700	358	BSRF1	278	61.5	ORF55	40.0	ORF55	42.0	Continued on following page
					BSLF1	218	46.9	ORF56	27.1	ORF56	26.5	
					BMLF1	874	38.8	ORF57	18.9	ORF57	19.9	
						438		K9				
								K10				
								K11				
ORF52	L	89773	88700	358	BMRF2	357	46.2	ORF58	21.8	ORF58	17.6	

Continued on following page

FIG. 4. Putative marmoset LCV Orip and comparison with EBV, rhesus LCV, and baboon LCV OriPs. (A) Scheme of the EBV, rhesus LCV, baboon LCV, and marmoset LCV Orip elements. Monomers of the family of repeats (FR) and the dyad symmetry element (DS) are represented as open boxes. The number and length of the monomers are indicated. Note that the marmoset LCV FR element contains two types of monomers, repeat A and repeat B. (B) Comparison of the consensus sequence of EBV, rhesus LCV, baboon LCV, and marmoset LCV repeat A and repeat B FR monomers. Consensus sequences for a given virus are shown as follows: a nucleotide conserved in all monomers is in boldface uppercase; a nucleotide conserved in more than 75% of the monomers is shown in lightface uppercase; and a nucleotide conserved in less than 75% of the monomers is shown in lower-case. X designates shorter monomers. The EBNA-1 binding site consensus sequence (TAGCATATGCTA) is boxed. (C) Comparison of the consensus sequence of EBV, rhesus LCV, and baboon LCV DS monomers. Consensus sequences are displayed with the criteria defined in B. No DS element could be identified in the marmoset LCV genome.

of the latent infection genes in rhesus, baboon, and marmoset LCVs (46%, 44%, and 36% amino acid homology to EBV EBNA-1, respectively). No equivalent to the DS element could be identified downstream of the marmoset LCV FR with the consensus sequence of either the EBNA-1 binding site or the marmoset LCV repeat A or B monomers. Whether this region represents a functional Orip important for episomal maintenance and initiating DNA replication during marmoset LCV latent infection remains to be determined.

Analysis of the marmoset LCV U_s sequence also showed that this repeat region was totally encompassed by an ORF, C6, which had no homology to other cell or viral genes. There was no ORF across the Orip region in the EBV or rhesus LCV genome; however, ORFs encompassing the lytic origin of replication, OriLyt, were present in the EBV, rhesus, and marmoset LCV genomes.

Marmoset LCV U_s : absence of vIL-10 and EBER homologues. The BCRF1 ORF in the EBV U_s had 84% amino acid homology and biological properties similar to those of the cellular interleukin-10 (cIL-10) and is referred to as viral IL-10 (vIL-10) (23, 64). A vIL-10 homologue is conserved in the rhesus LCV at the same location relative to the EBV genome, with 84% homology to EBV vIL-10 (48). In contrast, there was no ORF at the same position in the marmoset LCV genome, i.e., between the putative Orip and IR1. In addition, no ORF could be identified with homology to EBV, rhesus LCV, or cellular IL-10 anywhere in the marmoset LCV genome. Thus, the absence of an IL-10 homologue in this New World LCV and the high homology between vIL-10 and cIL-10 suggests

that LCVs may have acquired the IL-10 homologue from the cell genome at some point after the evolution of the marmoset LCV.

Similarly, there were no sequence homologues for the EBV EBERs in the marmoset LCV U_s . The baboon and rhesus LCV EBERs showed moderate sequence homology to EBV EBERs (40 to 70% nucleotide homology), and there are portions of nucleotide sequence that have been strongly conserved among human, baboon, and rhesus LCV EBERs corresponding to potentially important secondary stem-loop structures (22, 45). Searches with these consensus EBER nucleotide sequences against the complete marmoset LCV sequence also failed to detect any homologous sequences.

In order to test whether the marmoset LCV might express similar small, abundant RNAs from the U_s with no significant sequence homology to the EBV, rhesus LCV, or baboon LCV EBERs, RNA from marmoset LCV-infected cells was hybridized with a cosmid DNA probe containing the complete marmoset LCV U_s (Fig. 5, lower right panel). Several larger RNA species consistent with lytic infection transcripts were detected, but no small transcripts (<1 kb) were detected. As a control, RNA from rhesus LCV-infected cells was hybridized with a similar-sized cosmid DNA probe containing the rhesus LCV U_s . Again, several RNA transcripts greater than 1 kb in size were identified, but in addition, small (~0.3 kb) RNA transcripts were easily detected (Fig. 5, lower left panel). These were confirmed to be the rhesus LCV EBERs by hybridization with a more specific DNA probe (data not shown).

To test whether an EBER equivalent might be expressed from other regions of the marmoset LCV genome, total RNA was hybridized individually with four overlapping DNA probes covering the remainder of the marmoset LCV genome. However, there was no evidence for any small, abundant RNAs expressed in marmoset LCV-infected cells (data not shown). Thus, by both sequence and Northern blot analysis, there did not appear to be EBER homologues present in the marmoset LCV, suggesting that these EBV genes were also acquired later in LCV evolution.

Marmoset LCV IR1: C0, a leader sequence for transcripts derived by extensive splicing through the major internal repeat. EBV and rhesus LCV show a similar transcriptional pattern for the latent-infection EBNA genes, with a common promoter for all EBNA transcripts in the U_s or IR1 (16, 40, 54, 57). Since there were no obvious promoter elements immediately upstream of the marmoset LCV C5, C3, and ORF39 genes, i.e., the EBNA-2, EBNA-3, and EBNA-1 positional homologues, we asked whether these marmoset LCV transcripts might also show a common upstream promoter and be derived by alternative splicing.

5'RACE was performed with a primer at nucleotide 77 of C5. Ten independent RACE clones were sequenced, and all demonstrated a highly spliced 5' leader sequence reminiscent of the EBV and rhesus LCV EBNA-LP mRNAs present in all EBNA transcripts (40, 54). The marmoset LCV leader sequence included (i) 144 nucleotides immediately upstream of the C5 translational initiation codon (identified as exon E4 [E4]), (ii) two unique exons adjacent to IR1 (E2 and E3, analogous to the EBV EBNA-LP Y1 and Y2 exons, respectively), and (iii) repetitive elements derived by alternating exons spliced from sequential repeats within IR1 (E1b and E1a,

FIG. 5. Absence of EBERs in marmoset LCV-infected cells. Total RNA from EBV-negative BJAB cells, rhesus (Rh) LCV-infected LCL 8664 cells, or marmoset LCV-infected Cj0149 cells were separated by gel electrophoresis and visualized by ethidium bromide staining (top panels). Northern blots were hybridized with 32 P-labeled rhesus LCV cosmid CD1 DNA containing the entire U_S and IR1 and a portion of the U_L region of rhesus LCV (bottom left) or 32 P-labeled marmoset LCV cosmid E4 DNA containing the entire U_S and IR1 and a portion of the U_L region of marmoset LCV (bottom right). The positions of the rhesus LCV EBERs are shown by an arrow.

analogous to the EBV EBNA-LP W1 and W2 exons, respectively) (Fig. 2 and 6).

The longest 5' RACE products contained four pairs of E1 exons, suggesting that the promoter for C5 was further upstream in the U_S or IR1, as for human and Old World LCVs. The leader sequence contained a large open reading frame (Fig. 2), and it remains to be determined whether the 5' end of the leader sequence contains a codon capable of initiating translation of a protein product similar to EBNA-LP. This putative open reading frame had no sequence homology to EBNA-LP but was predicted to encode a nuclear protein highly phosphorylated on serine residues (Fig. 2) like EBNA-LP (68). Due to these extensive similarities, this putative EBNA-LP homologue was referred to as C0.

To determine whether the C0 leader sequence is shared by other putative marmoset LCV latent infection genes, RT-PCR of marmoset LCV-infected cell RNA was performed with a C0-specific primer in exon E3 combined with a C3- or ORF39-specific primer. Gel electrophoresis of C0-C3 and C0-ORF39 RT-PCR amplifications revealed two products from each reaction, differing by approximately 150 bp (Fig. 6). Sequencing of multiple clones for each of the C3 and ORF39 RT-PCR products revealed a similar leader sequence containing the unique C0 exons E3 and E4, spliced either directly to the C3 or ORF39 ORF or to an intermediate fifth exon consisting of 170 nucleotides located approximately 20 kb downstream of C0. This fifth exon (E5) was similar in location to the EBV Bam U exon present in the EBNA-1, -3A, -3B, and -3C transcripts

FIG. 6. Structure of the marmoset LCV C0, C5, C3, and ORF39 transcripts. Schemes for the marmoset LCV genome (A) and C5, C3, and ORF39 spliced mRNAs (B) are shown. Genome and mRNAs are not drawn to scale. Exons 1a and 1b of the common leader sequence (C0) are derived from the repeats in IR1, and C0 exons 2, 3, and 4 are derived from U_L . Alternative splicing of an additional fifth exon at nucleotides 102177 to 102007 for C3 and ORF39 transcripts is shown. RT-PCR amplification products for the respective transcripts are shown on the right. Locations of the primers used for the RT-PCR analysis are indicated by arrowheads. Control amplifications with all reagents except reverse transcriptase (-RT) and with water substituted for cDNA (W) are shown. The positions of size markers are shown in base pairs.

from EBV and rhesus LCV (3, 53). Thus, a common leader sequence for latent infection transcripts, the use of a common latent infection promoter in U_S and IR1, and the use of very similar splicing patterns to generate latent infection transcripts are strongly conserved in all LCVs.

Overview of complete marmoset LCV genome nucleotide sequence. The marmoset LCV genome sequence is now complete and provides a prototype sequence for New World LCVs. The genome sequence is 161,345 bp long, based on 7.5 copies of the 2,912-bp major IR and 35 copies of a GC-rich (75%), 328-bp terminal repeat. Overall, the marmoset LCV genome has a 50% GC content. Due to the order in which the marmoset LCV genome was cloned and sequenced, the complete marmoset LCV nucleotide sequence in the GenBank database is numbered in the opposite orientation relative to the EBV and rhesus LCV nucleotide coordinates (Fig. 7A) (6).

Whereas the rhesus LCV genome is predicted to have the same repertoire of ORFs as EBV (48), the marmoset LCV is significantly different from EBV. There are 73 predicted ORFs; 59 of these ORFs are homologous to genes found in all herpesviruses and most likely play a role in lytic viral infection (Fig. 7B, ORFs with bold outline). Six additional genes are homologous to genes present only in herpesviruses belonging to the LCV genus; ORFs 1, 6, 39, 43, 44, and 45 are homologous to EBV BALF1, BILF1, EBNA-1, BZLF1, BZLF2, and gp350, respectively. Overall, these 65 ORFs have 47% amino acid homology to EBV ORFs. In comparison, the rhesus LCV,

FIG. 7. Marmoset LCV genome and ORFs and homology to EBV ORFs. (A) Organization of the marmoset LCV genome. Major repeat regions IR1 (nt 142148 to 120197), IR2 (nt 117265 to 116619), IR3 (nt 69982 to 69841), IR4 (nt 2792 to 2469), and terminal repeats (TR, nt 161345 to 152171) are identified in the marmoset LCV genome. (B) Marmoset LCV ORFs and amino acid homology to EBV ORFs. The percent amino acid similarity is shown on the y axis. Putative latent, immediate-early, early, and late lytic ORFs are represented by black, dark grey, light grey, and white arrows, respectively. The ORFs are numbered from right to left, and the orientation of the ORFs is shown by the direction of the arrow. ORFs common to other herpesviruses are shown with a bold outline. The initiator codon for each ORF is positioned accurately, but the ORF size is not drawn to scale. (C) Marmoset LCV unique genes. ORFs are represented as defined for B.

a prototype for the Old World LCVs, has 75% overall amino acid homology to EBV (48).

Eight putative marmoset LCV genes have no amino acid sequence homology to known cell or viral genes and are identified as C0 through C7 (Fig. 7C). Five of these genes, C0, C1, C3, C5, and C7, are positional homologues for EBV latent infection genes EBNA-LP, LMP1, EBNA-3, EBNA-2, and LMP2, respectively, and are presumed to be marmoset LCV latent infection genes. Marmoset LCV C2 and C4 are positional homologues of, but have no sequence homology to, the EBV BILF2 and BHLF1, respectively. Marmoset LCV C6 is a putative ORF encoded across the Orip homologue in U_S, and there is no homologue present in either EBV or rhesus LCV.

Eleven genes found in EBV and rhesus LCV are not present in the marmoset LCV. These include the latent infection genes EBER and BARF0, the putative immunomodulators BCRF1 and BARF1, homologues to the cellular genes IL-10 and CSF1-R, respectively, the glycoprotein BDLF3, which is the most divergent glycoprotein in the Old World LCVs (48), and

other lytic genes whose contribution to EBV infection remains unclear (BLLF2, LF3, ECRF4, and BNLF2A/B).

DISCUSSION

Complete genome sequences of three different members of the LCV genus have now been determined, and these include LCVs infecting humans (EBV) (2), Old World primates (rhesus LCV) (48), and New World primates (LCV) (this report). In terms of the host, the Old and New World primates evolved approximately 25 million and 35 million years before humans, respectively (20, 59). If one takes the traditional approach of analyzing the nucleotide or amino acid homology for a given gene in EBV, rhesus LCV, and marmoset LCV, the homology will generally reflect this temporal relationship. For example, the overall amino acid homology of all ORFs in rhesus LCV is 75% with EBV ORFs (48), and the marmoset LCV ORFs are more distant, with only 43% amino acid homology to EBV ORFs.

Buried in this type of analysis is consideration for the background from which a given gene has evolved. For example, the DNA polymerase is strongly conserved in the rhesus and marmoset LCVs (94.8% and 73.5% amino acid homology to EBV, respectively), as might be expected since all herpesviruses replicate by a similar mechanism (reviewed in reference 52). The DNA polymerase is probably one of the oldest herpesvirus genes, so that the optimal fit at each peptide position has probably been achieved and there is little selective pressure for different amino acid substitutions. On the other hand, genes found only in herpesviruses of the LCV genus are likely to be younger genes and generally have less amino acid homology.

The LMP1 homologues from EBV, rhesus LCV, and marmoset LCV are the LCV-specific genes that have been studied in greatest detail to date. Marmoset LCV C1 is a positional and functional homologue of EBV LMP1 but has no significant sequence homology to its EBV counterpart (6). The rhesus LCV LMP1 has 32.4% overall amino acid homology to EBV LMP1, but virtually all of this homology is a result of the hydrophobic residues in the transmembrane domains (13). Despite this dramatic sequence divergence, the marmoset LCV C1, rhesus LCV LMP1, and EBV LMP1 all have similar secondary structures and similar functional mechanisms, i.e., interaction with tumor necrosis factor receptor-associated factors (TRAF) to activate NF- κ B (6, 13). The common secondary structure, i.e., short amino terminus, six hydrophobic transmembrane domains, and large carboxy terminus containing multiple TRAF interaction domains, suggests that the LMP1 homologues all evolved from some common progenitor. However, since these are relatively new genes, there has not been sufficient time for an optimal fit of amino acid residues to be achieved and for an optimized LMP1 sequence to be passed on to all primate LCVs, as for the DNA polymerase.

Biological pressures for membrane association and NF- κ B activation were probably the most stringent requirements for LMP1 homologues, and this is reflected in the high amino acid sequence homology of the transmembrane domains and TRAF interaction domains. There is less selective pressure on the intervening portions of the LMP1 homologues, resulting in a wider variety of acceptable sequences and thus more sequence divergence. One cannot determine from these types of data whether the DNA mutation rate during viral replication is any different between genes such as the DNA polymerase and the LMP1 homologues, but the lack of an optimal fit for the newer LMP1 homologues suggests that the chance of any single base change's being tolerated and perpetuated is higher for the LMP1 homologues than for the DNA polymerase homologues.

Another approach to comparative genetic analysis of the LCV genomes is to examine the genetic repertoire of the viruses as they have evolved. This approach may be more fruitful for linking viral genetics to viral pathogenesis. Genetic analyses in tissue culture systems have been very effective for defining viral genes essential for in vitro lytic infection replication and for LCV-induced B-cell transformation (reviewed in reference 28). However, after these studies, we are still left with 30 to 40% of the viral gene repertoire with no obvious role defined by in vitro systems, and one presumes that nonessential genes must be important for viral pathogenesis in vivo. The identical gene repertoires of the rhesus LCV and EBV are therefore extremely valuable because this provides an experi-

TABLE 2. Classification of LCV-specific genes based on evolutionary analysis

LCV gene class	Sequence	Gene	Description
Ancestral (present in all LCVs)	Conserved	BLLF1-gp350	
		BZLF2-gp42	
		BILF1-gp64	
		BZLF1	Transactivator
		EBNA-1	Latent gene
		BHRF1	Bcl-2 homologue
		BALF1	Bcl-2 homologue
		LF1	Unknown function
		LF2	Unknown function
	Divergent	BILF2-gp78	
		LMP2	Latent gene
		EBNA-LP	Latent gene
		EBNA-2	Latent gene
		EBNA-3	Latent gene
		LMP1	Latent gene
		BHLF1	Unknown function
Acquired (present only in Old World LCVs)		BCRF1 vIL-10	Immunomodulator
		BARF1 CSF-1R	Immunomodulator
		BDLF3-gp150	
		EBERs	Latent gene
		BARF0	Latent gene
		BLLF2	Unknown function
		LF3	Unknown function
		ECRF4	Unknown function
		BNLF2s	Unknown function

mental animal model system in which the in vivo role in viral pathogenesis of these nonessential viral genes can be studied (48). The different genetic repertoire of the marmoset LCV provides insight into LCV evolution by differentiating those viral genes that were present in the earliest LCVs versus those that were acquired later in LCV evolution.

With the marmoset LCV as a defining point, we can now propose three evolutionary classes of LCV genes (Table 2). (i) Universal herpesvirus genes, common to all herpesviruses, including genes important for viral replication and viral structure. These represent the vast majority of genes in any given LCV (68 to 80%) and are generally the most well conserved through LCV evolution (49% to 83% amino acid homology for marmoset LCV and rhesus LCV genes, respectively, compared with the EBV genes). (ii) Ancestral LCV-specific genes that are specific to the LCV genus, i.e., not found in other herpesviruses and present in New World LCVs, the earliest LCV known to date. (iii) Acquired LCV-specific genes that have evolved later in the LCV genus and are present in the Old World and human LCVs but absent from the New World LCVs. The ancestral LCV-specific genes can be further divided into two groups, conserved genes, i.e., those with amino acid homology to EBV, and divergent genes, i.e., genes that are positional homologues to EBV genes but show no significant homology to these genes.

The ancestral LCV-specific genes are by definition present in all LCVs. Therefore, it is not surprising to find viral genes involved in intrinsic LCV properties such as B-cell tropism, B-cell transformation, and the ability to reactivate in B cells. As shown in Table 2, this group includes two viral glycoproteins, BLLF1-gp350 and BZLF2-gp42, that are important for

B-cell infection by interacting with the EBV receptor CR2 and the coreceptor HLA class II (31, 38, 58, 62). Likewise, homologues for all the EBV latent genes essential for B-cell transformation, namely EBNA-LP, EBNA-2, EBNA-3, EBNA-1, and LMP1 (11, 19, 26, 27, 30, 34, 63), are present in all LCVs (reviewed in reference 66), and the transcriptional organization and complex splicing patterns required for expression of the latent infection nuclear proteins also appear to be an early evolutionary event for LCV. BZLF1, the lytic transactivator sufficient to induce the transition from latency to lytic infection (18), is also present in all three LCVs. Two bcl2 homologues captured by LCV (BHRF1 and BALF1), two other glycoproteins (BILF1-gp64 and BILF2-gp78), and three genes (LF1, LF2, and BHLF1) with no known function are also present in this group. Discovery of other LCVs may result in refinement of this list, but conceptually the repertoire of genes included in this ancestral group is consistent with biological properties that may be expected for a progenitor of the LCV genus.

The acquired LCV-specific genes are by definition present in Old World LCVs but absent from the New World LCVs. Most of these genes, namely BCRF1, BARF1, BDLF3, EBERs, BARF0, and LF3, have been shown to be nonessential for EBV replication and B-cell growth transformation in vitro (4, 10, 39, 51, 60, 61). These nonessential genes might be beneficial to LCV infection in vivo by facilitating either the invasion of the host during acute infection, the evasion of the host immune response during persistent infection, or the reactivation of viral infection to produce infectious progeny capable of infecting additional hosts. Two of the acquired LCV-specific genes are cellular homologues with immunomodulatory properties, BCRF1 (vIL-10) and BARF1 (CSF-1R). vIL-10 is thought to indirectly inhibit the initial natural killer (NK) and T-cell cytotoxic response to EBV infection via its negative effect on macrophages and dendritic cells (36). BARF1 inhibits CSF-1-induced gamma interferon (IFN- γ) release by activated monocytes in vitro and therefore might play a role in evading an early host response to EBV infection (10).

In addition to these two immunomodulators, the EBV-encoded RNAs, or EBERs, can inhibit protein kinase R activity in vitro and may be an important immunomodulator in vivo by inhibiting the negative effect of IFN- γ on viral protein synthesis (55). Another acquired gene, BDLF3, encodes a glycoprotein that has no effect on B-cell infection but may detrimentally affect the ability of EBV to infect epithelial cells (4). The duplication of the EBNA-3 genes also appears to be an acquired event in LCV evolution, as originally suggested by the remote sequence similarity and exon-intron structure of EBNA-3A, -3B, and -3C (21, 24, 41, 42). It remains to be determined whether the marmoset LCV gene C3 is functionally similar to the EBNA-3s by interacting with CBF-1/RBP-J κ and modulating CBF-1-mediated EBNA-2 transcriptional transactivation (25, 50, 65). It is interesting that another acquired gene, BARF0, and other potential ORFs in the *Bam*HI rightward transcripts have recently been reported to interact with NOTCH4 and CBF-1 (29, 56), suggesting an increasing importance for regulating CBF-1-dependent transcription during LCV evolution.

The acquisition of these additional accessory genes by Old World LCVs might be a consequence of coevolution with the host. LCV may have acquired these additional accessory genes

in response to host evolution in order to maintain its biological phenotype. Alternatively, the acquisition of these accessory genes by Old World LCVs and EBV may have been the initial driving event resulting in some new or modified biological properties not present in the New World LCVs. Testing this hypothesis will require more careful study of marmoset LCV infection and comparison with EBV and Old World LCV. For example, preliminary analyses with LCV-specific serological assays showed a much lower prevalence of marmoset LCV infection (40 to 60%) than rhesus LCV infection (95%) in captive colonies (C. Quink, A. Carville, and F. Wang, unpublished data). Whether this represents a fundamental difference in the biological behavior of marmoset LCV versus rhesus LCV or a consequence of domestic housing and breeding remains to be determined.

The origin of LCV remains a mystery. The rhadinoviruses are the most closely related herpesviruses biologically and genetically (for a review, see reference 37), and the rhadinoviruses have a much broader host range, suggesting that they are much older than LCVs. Thus, it is not unreasonable to hypothesize that LCV may have evolved from a branch of the rhadinoviruses. The genomes are similar in organization, so that the universal herpesvirus genes are typically encoded in the same colinear order and in blocks arising at similar locations throughout the LCV and rhadinovirus genomes (1). The LCV- and rhadinovirus-specific genes are also interspersed in the genomes within divergent blocks that arise at similar relative locations in the LCV and rhadinovirus genomes.

The marmoset LCV C7 may provide more specific evidence for a transition from rhadinoviruses to LCV. C7 is more similar to Kaposi's sarcoma-associated herpesvirus K15 by the presence of phosphotyrosine residues in the larger carboxy terminus, versus the amino terminus of EBV LMP2A (8, 17). However, C7 is more like the EBV LMP2, with transcription across the terminal repeats. The location of the 5'-most C7 exon within a C1 intron is an intermediate position between the LMP2A and LMP2B first exons, suggesting an ongoing evolutionary process at this locus. More definitive linkage between the rhadinovirus and LCV genera may require identification of earlier ancestors in the LCV genus potentially infecting other primates or mammals.

ACKNOWLEDGMENTS

We thank members of the Massachusetts General Hospital DNA Sequencing Core for performance of high-throughput DNA sequencing and Elliott Kieff for valuable advice.

This work was supported by a grant from the National Cancer Institute (CA89172).

REFERENCES

1. Albrecht, J. C., J. Nicholas, D. Biller, K. R. Cameron, B. Biesinger, C. Newman, S. Wittmann, M. A. Craxton, H. Coleman, B. Fleckenstein, et al. 1992. Primary structure of the herpesvirus saimiri genome. *J. Virol.* **66**:5047-5058.
2. Baer, R., A. T. Bankier, M. D. Biggin, P. L. Deininger, P. J. Farrell, T. J. Gibson, G. Hatfull, G. S. Hudson, S. C. Satchwell, C. Seguin, et al. 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature* **310**:207-211.
3. Bodescot, M., and M. Perricaudet. 1986. Epstein-Barr virus mRNAs produced by alternative splicing. *Nucleic Acids Res.* **14**:7103-7114.
4. Borza, C. M., and L. M. Hutt-Fletcher. 1998. Epstein-Barr virus recombinant lacking expression of glycoprotein gp150 infects B cells normally but is enhanced for infection of epithelial cells. *J. Virol.* **72**:7577-7582.
5. Burkhardt, A. L., J. B. Bolen, E. Kieff, and R. Longnecker. 1992. An Epstein-

- Barr virus transformation-associated membrane protein interacts with src family tyrosine kinases. *J. Virol.* **66**:5161–5167.
6. **Cho, Y., J. Ramer, P. Rivaller, C. Quink, R. L. Garber, D. R. Beier, and F. Wang.** 2001. An Epstein-Barr-related herpesvirus from marmoset lymphomas. *Proc. Natl. Acad. Sci. USA* **98**:1224–1229.
 7. **Cho, Y. G., A. V. Gordadze, P. D. Ling, and F. Wang.** 1999. Evolution of two types of rhesus lymphocryptovirus similar to type 1 and type 2 Epstein-Barr virus. *J. Virol.* **73**:9206–9212.
 8. **Choi, J. K., B. S. Lee, S. N. Shim, M. Li, and J. U. Jung.** 2000. Identification of the novel K15 gene at the rightmost end of the Kaposi's sarcoma-associated herpesvirus genome. *J. Virol.* **74**:436–446.
 9. **Cohen, J. I., and E. Kieff.** 1991. An Epstein-Barr virus nuclear protein 2 domain essential for transformation is a direct transcriptional activator. *J. Virol.* **65**:5880–5885.
 10. **Cohen, J. I., and K. Lekstrom.** 1999. Epstein-Barr virus BARF1 protein is dispensable for B-cell transformation and inhibits alpha interferon secretion from mononuclear cells. *J. Virol.* **73**:7627–7632.
 11. **Cohen, J. I., F. Wang, J. Mannick, and E. Kieff.** 1989. Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation. *Proc. Natl. Acad. Sci. USA* **86**:9558–9562.
 12. **Dykstra, M. L., R. Longnecker, and S. K. Pierce.** 2001. Epstein-Barr virus coopts lipid rafts to block the signaling and antigen transport functions of the BCR. *Immunity* **14**:57–67.
 13. **Franken, M., O. Devergne, M. Rosenzweig, B. Annis, E. Kieff, and F. Wang.** 1996. Comparative analysis identifies conserved tumor necrosis factor receptor-associated factor 3 binding sites in the human and simian Epstein-Barr virus oncogene LMP1. *J. Virol.* **70**:7819–7826.
 14. **Fruehling, S., and R. Longnecker.** 1997. The immunoreceptor tyrosine-based activation motif of Epstein-Barr virus LMP2A is essential for blocking BCR-mediated signal transduction. *Virology* **235**:241–251.
 15. **Fruehling, S., R. Swart, K. M. Dolwick, E. Kremmer, and R. Longnecker.** 1998. Tyrosine 112 of latent membrane protein 2A is essential for protein tyrosine kinase loading and regulation of Epstein-Barr virus latency. *J. Virol.* **72**:7796–7806.
 16. **Fuentes-Panana, E. M., S. Swaminathan, and P. D. Ling.** 1999. Transcriptional activation signals found in the Epstein-Barr virus (EBV) latency C promoter are conserved in the latency C promoter sequences from baboon and rhesus monkey EBV-like lymphocryptoviruses (cercopithecine herpesviruses 12 and 15). *J. Virol.* **73**:826–833.
 17. **Glenn, M., L. Rainbow, F. Aurad, A. Davison, and T. F. Schulz.** 1999. Identification of a spliced gene from Kaposi's sarcoma-associated herpesvirus encoding a protein with similarities to latent membrane proteins 1 and 2A of Epstein-Barr virus. *J. Virol.* **73**:6953–6963.
 18. **Grogan, E., H. Jensen, J. Countryman, L. Heston, L. Gradoville, and G. Miller.** 1987. Transfection of a rearranged viral DNA fragment, WZhet, stably converts latent Epstein-Barr virus infection to productive infection in lymphoid cells. *Proc. Natl. Acad. Sci. USA* **84**:1332–1336.
 19. **Hammerschmidt, W., and B. Sugden.** 1989. Genetic analysis of immortalizing functions of Epstein-Barr virus in human B lymphocytes. *Nature* **340**:393–397.
 20. **Hasegawa, M.** 1990. Phylogeny and molecular evolution in primates. *Jpn. J. Genet.* **65**:243–266.
 21. **Hennesy, K., F. Wang, E. W. Bushman, and E. Kieff.** 1986. Definitive identification of a member of the Epstein-Barr virus nuclear protein 3 family. *Proc. Natl. Acad. Sci. USA* **83**:5693–5697.
 22. **Howe, J. G., and M. D. Shu.** 1988. Isolation and characterization of the genes for two small RNAs of herpesvirus papio and their comparison with Epstein-Barr virus-encoded EBER RNAs. *J. Virol.* **62**:2790–2798.
 23. **Hsu, D. H., R. de Waal Malefyt, D. F. Fiorentino, M. N. Dang, P. Vieira, J. de Vries, H. Spits, T. R. Mosmann, and K. W. Moore.** 1990. Expression of interleukin-10 activity by Epstein-Barr virus protein BCRF1. *Science* **250**:830–832.
 24. **Joab, I., D. T. Rowe, M. Bodescot, J. C. Nicolas, P. J. Farrell, and M. Perricaudet.** 1987. Mapping of the gene coding for Epstein-Barr virus-determined nuclear antigen EBNA3 and its transient overexpression in a human cell line by with an adenovirus expression vector. *J. Virol.* **61**:3340–3344.
 25. **Johannsen, E., C. L. Miller, S. R. Grossman, and E. Kieff.** 1996. EBNA-2 and EBNA-3C extensively and mutually exclusively associate with RBPJ κ in Epstein-Barr virus-transformed B lymphocytes. *J. Virol.* **70**:4179–4183.
 26. **Kaye, K. M., K. M. Izumi, and E. Kieff.** 1993. Epstein-Barr virus latent membrane protein 1 is essential for B-lymphocyte growth transformation. *Proc. Natl. Acad. Sci. USA* **90**:9150–9154.
 27. **Kempkes, B., D. Pich, R. Zeidler, B. Sugden, and W. Hammerschmidt.** 1995. immortalization of human B lymphocytes by a plasmid containing 71 kilobase pairs of Epstein-Barr virus DNA. *J. Virol.* **69**:231–238.
 28. **Kieff, E., and A. B. Rickinson.** 2001. Epstein-Barr virus and its replication, p. 2511–2573. *In* D. Knipe, P. Howley, et al. (ed.), *Fields virology*, 4th ed., vol. 2. Raven Press, Philadelphia, Pa.
 29. **Kusano, S., and N. Raab-Traub.** 2001. An Epstein-Barr virus protein interacts with Notch. *J. Virol.* **75**:384–395.
 30. **Lee, M. A., M. E. Diamond, and J. L. Yates.** 1999. Genetic evidence that EBNA-1 is needed for efficient, stable latent infection by Epstein-Barr virus. *J. Virol.* **73**:2974–2982.
 31. **Li, Q., M. K. Spriggs, S. Kovats, S. M. Turk, M. R. Comeau, B. Nepom, and L. M. Hutt-Fletcher.** 1997. Epstein-Barr virus uses HLA class II as a cofactor for infection of B lymphocytes. *J. Virol.* **71**:4657–4662.
 32. **Loeb, D. D., N. S. Sung, R. L. Pesano, C. J. Sexton, C. D. Hutchison, and J. S. Pagano.** 1990. Plasmid origin of replication of herpesvirus papio: DNA sequence and enhancer function. *J. Virol.* **64**:2876–2883.
 33. **Longnecker, R., B. Druker, T. M. Roberts, and E. Kieff.** 1991. An Epstein-Barr virus protein associated with cell growth transformation interacts with a tyrosine kinase. *J. Virol.* **65**:3681–3692.
 34. **Mannick, J. B., J. I. Cohen, M. Birkenbach, A. Marchini, and E. Kieff.** 1991. The Epstein-Barr virus nuclear protein encoded by the leader of the EBNA RNAs is important in B-lymphocyte transformation. *J. Virol.* **65**:6826–6837.
 35. **Moghaddam, A., M. Rosenzweig, D. Lee-Parritz, B. Annis, R. P. Johnson, and F. Wang.** 1997. An animal model for acute and persistent Epstein-Barr virus infection. *Science* **276**:2030–2033.
 36. **Moore, K. W., R. de Waal Malefyt, R. L. Coffman, and A. O'Garra.** 2001. Interleukin-10 and the interleukin-10 receptor. *Annu. Rev. Immunol.* **19**:683–765.
 37. **Moore, P. S., and Y. Chang.** 2001. Kaposi's sarcoma-associated herpesvirus, p. 2803–2833. *In* D. Knipe, P. Howley, et al. (ed.), *Fields virology*, 4th ed., vol. 2. Raven Press, Philadelphia, Pa.
 38. **Nemerow, G. R., C. Mold, V. K. Schwend, V. Tollefson, and N. R. Cooper.** 1987. Identification of gp350 as the viral glycoprotein mediating attachment of Epstein-Barr virus (EBV) to the EBV/C3d receptor of B cells: sequence homology of gp350 and C3 complement fragment C3d. *J. Virol.* **61**:1416–1420.
 39. **Parker, B. D., A. Bankier, S. Satchwell, B. Barrell, and P. J. Farrell.** 1990. Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region. *Virology* **179**:339–346.
 40. **Peng, R., A. V. Gordadze, E. M. Fuentes Panana, F. Wang, J. Zong, G. S. Hayward, J. Tan, and P. D. Ling.** 2000. Sequence and functional analysis of EBNA-LP and EBNA2 proteins from nonhuman primate lymphocryptoviruses. *J. Virol.* **74**:379–389.
 41. **Petti, L., and E. Kieff.** 1988. A sixth Epstein-Barr virus nuclear protein (EBNA3B) is expressed in latently infected growth-transformed lymphocytes. *J. Virol.* **62**:2173–2178.
 42. **Petti, L., J. Sample, F. Wang, and E. Kieff.** 1988. A fifth Epstein-Barr virus nuclear protein (EBNA3C) is expressed in latently infected growth-transformed lymphocytes. *J. Virol.* **62**:1330–1338.
 43. **Ramer, J. C., R. L. Garber, K. E. Steele, J. F. Boyson, C. O'Rourke, and J. A. Thomson.** 2000. Fatal lymphoproliferative disease associated with a novel gammaherpesvirus in a captive population of common marmosets. *Comp. Med.* **50**:59–68.
 44. **Rangan, S. R., L. N. Martin, B. E. Bozelka, N. Wang, and B. J. Gormus.** 1986. Epstein-Barr virus-related herpesvirus from a rhesus monkey (*Macaca mulatta*) with malignant lymphoma. *Int. J. Cancer* **38**:425–432.
 45. **Rao, P., H. Jiang, and F. Wang.** 2000. Cloning of the rhesus lymphocryptovirus viral capsid antigen and Epstein-Barr virus-encoded small RNA homologues and use in diagnosis of acute and persistent infections. *J. Clin. Microbiol.* **38**:3219–3225.
 46. **Reisman, D., J. Yates, and B. Sugden.** 1985. A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two *cis*-acting components. *Mol. Cell. Biol.* **5**:1822–1832.
 47. **Rickinson, A. B., and E. Kieff.** 2001. Epstein-Barr virus, p. 2575–2627. *In* D. Knipe, P. Howley, et al. (ed.), *Fields virology*, 4th ed., vol. 2. Raven Press, Philadelphia, Pa.
 48. **Rivaller, P., H. Jiang, Y. G. Cho, C. Quink, and F. Wang.** 2002. Complete nucleotide sequence of the rhesus lymphocryptovirus: genetic validation for an Epstein-Barr virus animal model. *J. Virol.* **76**:421–426.
 49. **Rivaller, P., C. Quink, and F. Wang.** 1999. Strong selective pressure for evolution of an Epstein-Barr virus LMP2B homologue in the rhesus lymphocryptovirus. *J. Virol.* **73**:8867–8872.
 50. **Robertson, E. S., J. Lin, and E. Kieff.** 1996. The amino-terminal domains of Epstein-Barr virus nuclear proteins 3A, 3B, and 3C interact with RBPJ κ . *J. Virol.* **70**:3068–3074.
 51. **Robertson, E. S., B. Tomkinson, and E. Kieff.** 1994. An Epstein-Barr virus with a 58-kilobase-pair deletion that includes BARF0 transforms B lymphocytes *in vitro*. *J. Virol.* **68**:1449–1458.
 52. **Roizman, B., and P. E. Pellet.** 2001. The family herpesviridae: a brief introduction, p. 2381–2397. *In* D. Knipe and P. Howley (ed.), *Fields virology*, 4th ed., vol. 2. Raven Press, Philadelphia, Pa.
 53. **Ruf, I. K., A. Moghaddam, F. Wang, and J. Sample.** 1999. Mechanisms that regulate Epstein-Barr virus EBNA-1 gene transcription during restricted latency are conserved among lymphocryptoviruses of Old World primates. *J. Virol.* **73**:1980–1989.
 54. **Sample, J., M. Hummel, D. Braun, M. Birkenbach, and E. Kieff.** 1986. Nucleotide sequences of mRNAs encoding Epstein-Barr virus nuclear proteins: a probable transcriptional initiation site. *Proc. Natl. Acad. Sci. USA* **83**:5096–5100.
 55. **Sharp, T. V., M. Schwemmle, I. Jeffrey, K. Laing, H. Mellor, C. G. Proud, K.**

- Hilse, and M. J. Clemens. 1993. Comparative analysis of the regulation of the interferon-inducible protein kinase PKR by Epstein-Barr virus RNAs EBER-1 and EBER-2 and adenovirus VAI RNA. *Nucleic Acids Res.* **21**: 4483–4490.
56. Smith, P. R., O. de Jesus, D. Turner, M. Hollyoake, C. E. Karstegl, B. E. Griffin, L. Karran, Y. Wang, S. D. Hayward, and P. J. Farrell. 2000. Structure and coding content of CST (BART) family RNAs of Epstein-Barr virus. *J. Virol.* **74**:3082–3092.
 57. Speck, S. H., A. Pfitzner, and J. L. Strominger. 1986. An Epstein-Barr virus transcript from a latently infected, growth-transformed B-cell line encodes a highly repetitive polypeptide. *Proc. Natl. Acad. Sci. USA* **83**:9298–9302.
 58. Spriggs, M. K., R. J. Armitage, M. R. Comeau, L. Strockbine, T. Farrah, B. Macduff, D. Ulrich, M. R. Alderson, J. Mullberg, and J. I. Cohen. 1996. The extracellular domain of the Epstein-Barr virus BZLF2 protein binds the HLA-DR beta chain and inhibits antigen presentation. *J. Virol.* **70**:5557–5563.
 59. Stewart, C. B., and T. R. Disotell. 1998. Primate evolution — in and out of Africa. *Curr Biol.* **8**:R582–R588.
 60. Swaminathan, S., R. Hesselton, J. Sullivan, and E. Kieff. 1993. Epstein-Barr virus recombinants with specifically mutated BCRF1 genes. *J. Virol.* **67**: 7406–7413.
 61. Swaminathan, S., B. Tomkinson, and E. Kieff. 1991. Recombinant Epstein-Barr virus with small RNA (EBER) genes deleted transforms lymphocytes and replicates in vitro. *Proc. Natl. Acad. Sci. USA* **88**:1546–1550.
 62. Tanner, J., J. Weis, D. Fearon, Y. Whang, and E. Kieff. 1987. Epstein-Barr virus gp350/220 binding to the B lymphocyte C3d receptor mediates adsorption, capping, and endocytosis. *Cell* **50**:203–213.
 63. Tomkinson, B., E. Robertson, and E. Kieff. 1993. Epstein-Barr virus nuclear proteins EBNA-3A and EBNA-3C are essential for B-lymphocyte growth transformation. *J. Virol.* **67**:2014–2025.
 64. Vieira, P., R. de Waal-Malefyt, M. N. Dang, K. E. Johnson, R. Kastelein, D. F. Fiorentino, J. E. de Vries, M. G. Roncarolo, T. R. Mosmann, and K. W. Moore. 1991. Isolation and expression of human cytokine synthesis inhibitory factor cDNA clones: homology to Epstein-Barr virus open reading frame BCRF1. *Proc. Natl. Acad. Sci. USA* **88**:1172–1176.
 65. Waltzer, L., M. Perricaudet, A. Sergeant, and E. Manet. 1996. Epstein-Barr virus EBNA3A and EBNA3C proteins both repress RBP-J κ -EBNA2-activated transcription by inhibiting the binding of RBP-J κ to DNA. *J. Virol.* **70**:5909–5915.
 66. Wang, F., P. Rivailier, P. Rao, and Y. Cho. 2001. Simian homologues of Epstein-Barr virus. *Phil. Trans. R. Soc. Lond. B* **356**:489–497.
 67. Yalamanchili, R., S. Harada, and E. Kieff. 1996. The N-terminal half of EBNA2, except for seven prolines, is not essential for primary B-lymphocyte growth transformation. *J. Virol.* **70**:2468–2473.
 68. Yokoyama, A., M. Tanaka, G. Matsuda, K. Kato, M. Kanamori, H. Kawasaki, H. Hirano, I. Kitabayashi, M. Ohki, K. Hirai, and Y. Kawaguchi. 2001. Identification of major phosphorylation sites of Epstein-Barr virus nuclear antigen leader protein (EBNA-LP): ability of EBNA-LP to induce latent membrane protein 1 cooperatively with EBNA-2 is regulated by phosphorylation. *J. Virol.* **75**:5119–5128.
 69. Zimmer, U., H. K. Adldinger, G. M. Lenoir, M. Vuillaume, M. V. Knebel-Doeberitz, G. Laux, C. Desgranges, P. Wittmann, U. K. Freese, U. Schneider, et al. 1986. Geographical prevalence of two types of Epstein-Barr virus. *Virology* **154**:56–66.