JOURNAL OF VIROLOGY, Aug. 1992, p. 5047–5058
0022-538X/92/085047-12$02.00/0

Vol. 66, No. 8

Primary Structure of the Herpesvirus Saimiri Genome†

JENS-CHRISTIAN ALBRECHT,[1]* JOHN NICHOLAS,[2]‡ DORIS BILLER,[1] KEITH R. CAMERON,[2] BRIGITTE BIESINGER,[1] CAROL NEWMAN,[2] SABINE WITTMANN,[1] MOLLY A. CRAXTON,[2]§ HEATHER COLEMAN,[2] BERNHARD FLECKENSTEIN,[1] AND ROBERT W. HONESS[2]

Institut für Klinische und Molekulare Virologie, Friedrich-Alexander Universität, Loschgestrasse 7, D-8520 Erlangen, Germany,[1] and National Institute for Medical Research, Mill Hill, London NW7 1AA, United Kingdom[2]

Received 16 March 1992/Accepted 1 May 1992

This report describes the complete nucleotide sequence of the genome of herpesvirus saimiri, the prototype of gammaherpesvirus subgroup 2 (rhadinoviruses). The unique low-G+C-content DNA region has 112,930 bp with an average base composition of 34.5% G+C and is flanked by about 35 noncoding high-G+C-content DNA repeats of 1,444 bp (70.8% G+C) in tandem orientation. We identified 76 major open reading frames and a set of seven U-RNA genes for a total of 83 potential genes. The genes are closely arranged, with only a few regions of sizable noncoding sequences. For 60 of the predicted proteins, homologous sequences are found in other herpesviruses. Genes conserved between herpesvirus saimiri and Epstein-Barr virus (gammaherpesvirus subgroup 1) show that their genomes are generally collinear, although conserved gene blocks are separated by unique genes that appear to determine the particular phenotype of these viruses. Several deduced protein sequences of herpesvirus saimiri without counterparts in most of the other sequenced herpesviruses exhibited significant homology with cellular proteins of known function. These include thymidylate synthase, dihydrofolate reductase, complement control proteins, the cell surface antigen CD59, cyclins, and G protein-coupled receptors. Searching for functional protein motifs revealed that the virus may encode a cytosine-specific methylase and a tyrosine-specific protein kinase. Several herpesvirus saimiri genes are potential candidates to cooperate with the gene for saimiri transformation-associated protein of subgroup A (STP-A) in T-lymphocyte growth stimulation.

Herpesvirus saimiri is a virus of squirrel monkeys (*Saimiri sciureus*) that can be consistently isolated from peripheral mononuclear blood cells of healthy animals (36). The virus causes fulminant T-cell lymphomas in New World primates other than its natural hosts (37). Herpesvirus saimiri is also capable of transforming simian (32, 100, 105) and human (12) T lymphocytes to continuous growth in vitro. The known mammalian herpesviruses are usually grouped into three subfamilies: the large, heterogeneous alphaherpesvirus group, which includes herpes simplex virus and varicella-zoster virus; the betaherpesvirus group, represented by human cytomegalovirus (HCMV) and, on the basis of available sequence data, human herpesvirus 6; and the gammaherpesvirus group, which comprises herpesviruses that replicate and persist in lymphocytes, inducing lymphoproliferation (94, 95). While the well-characterized B-lymphotropic human Epstein-Barr virus (EBV) represents gammaherpesvirus subgroup 1, herpesvirus saimiri is the prototype of subgroup 2 gammaherpesviruses (rhadinoviruses) (44, 94, 95). Other subgroup 2 gammaherpesviruses are herpesvirus ateles and herpesvirus aotus, which are simian isolates; herpesvirus sylvilagus, which infects cottontail rabbits (73); bovine herpesvirus 4 (102); alcelaphine herpesvirus 1 (15); and murine herpesvirus 68 (35). All of these viruses share a characteristic genome structure with a unique internal low-G+C-content DNA segment (L-DNA) of about 110 kbp that is flanked by multiple tandem repetitions of high-G+C-content DNA (H-DNA). Each repeat unit of herpesvirus saimiri has 1,444 bp and a G+C content of 70.8% (9, 101). The oncogenicity and T-cell-transforming capability of herpesvirus saimiri were assigned to a left-terminal L-DNA region that is heterogeneous in different herpesvirus saimiri wild-type isolates (30, 56, 74). Nucleotide sequencing of this region revealed the saimiri transformation-associated protein (STP) reading frames, which encode a novel family of herpesvirus oncogene products (13, 50, 51, 79). A number of other genes of herpesvirus saimiri whose primary structures were determined are not conserved in the herpesvirus family; these include U-RNA genes (4, 13, 65, 78, 110), a dihydrofolate reductase (DHFR) gene (108), a thymidylate synthase (TS) gene (14, 45), a gene for putative complement-regulating viral glycoproteins (3), a gene for an immediate-early protein with sequence similarity to the minor lymphocyte-stimulatory (Mls) proteins of mice (85, 107), and two reading frames encoding homologs of cyclins and G protein-coupled receptors (82). Earlier sequencing studies showed that the genomes of herpesvirus saimiri and EBV are predominantly collinear (2, 16, 38, 39, 46, 81, 83, 84) and that homologous genes are in approximately equivalent locations and the same relative orientations. In this article, we present the complete DNA sequence of the coding L-DNA of herpesvirus saimiri and discuss the structural and biological relevance of the majority of the genes.

* Corresponding author.

† This article is dedicated to the memory of Robert W. Honess, who initiated this endeavor and died before its completion.

‡ Present address: Johns Hopkins Oncology Center, Baltimore, MD 21231.

§ Present address: MRC Centre, Cambridge CB2 2QH, United Kingdom.

MATERIALS AND METHODS

Virus and cell culture and plasmid and cosmid clones. Herpesvirus saimiri strain 11 (36), a virus of subgroup A (74), was propagated in owl monkey kidney (OMK 637) cells and

cloned into plasmid pACYC184 or pWD7 or cosmid pJC81 or λ Charon 4A (55). For sequencing purposes, restriction fragments or derived nested deletions were cloned into phagemid pBluescript KS+, pBluescript KS−, pBS+, or pBS− (Stratagene, La Jolla, Calif.) or into plasmid pUC18 or pUC19, or random fragments of sonicated herpesvirus saimiri DNA were cloned into the *Sma*I site of phage M13mp18 or M13mp19 as described by Bankier and Barrell (8). Recombinant DNA was purified through CsCl density gradients or by anion-exchange chromatography (Qiagen Inc.).

DNA sequencing. Nucleotide sequencing of cloned fragments of herpesvirus saimiri DNA was done either by direct sequencing of double-stranded DNA using synthetic oligonucleotide primers to "walk" along the fragments (*Kpn*I-B, *Kpn*I-F, *Kpn*I-C, *Eco*RI-B, and part of *Eco*RI-D [Institut für Klinische und Molekulare Virologie, Friedrich-Alexander Universität, Erlangen, Germany]) or by sequencing single-stranded M13 clones (*Kpn*I-E, *Eco*RI-C, *Eco*RI-E, *Eco*RI-J, *Eco*RI-H, and part of *Eco*RI-D [National Institute for Medical Research, London, United Kingdom]) by the dideoxynucleotide chain-termination method. Junctions between the *Kpn*I and *Eco*RI clones were confirmed by sequencing overlapping clones. Most nucleotides were determined on both strands by several independent reactions, and each autoradiogram was read at least three times. Oligonucleotide primers were prepared in a Biosearch Inc. Cyclone DNA synthesizer or purchased commercially.

Computer analysis of DNA and protein sequences. The sequence data were assembled and analyzed with the Genetics Computer Group, Inc., Sequence Analysis Software Package (version 7) (33) implemented on a MicroVAX 3500 computer (Digital Equipment Corporation) or as described previously (16). Details of the assembly and analysis of the rightmost 43,658 bp of herpesvirus saimiri L-DNA have been described elsewhere (81). The data bases used were GenBank (release 69; 55, 631 entries and 71,947,426 bases), PIR (release 29; 31,895 entries and 9,091,049 residues), SwissProt (release 19; 21,795 entries and 7,173,785 amino acids), and PROSITE (release 7.1; 509-protein consensus pattern). The PROSITE *Dictionary of Protein Sites and Patterns* by Amos Bairoch of the University of Geneva, Geneva, Switzerland, is distributed by the European Molecular Biology Laboratory (EMBL).

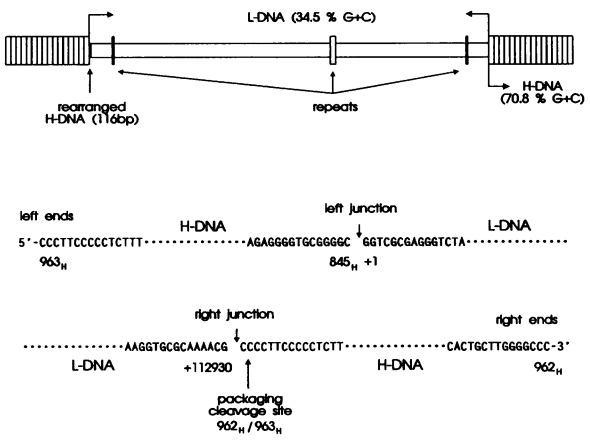

FIG. 1. Genome structure of herpesvirus saimiri. The genome of herpesvirus saimiri consists of a light component (L-DNA; 34.5% G+C) that is flanked by tandemly repeated heavy components (H-DNA; 70.8% G+C). The first and last base pairs of L-DNA (positions +1 and +112930, respectively) are the first base pairs to digress from the standard H-DNA repeat unit (9, 79). By this definition, the first 116 bp of L-DNA are composed of rearranged H-DNA. Positions +1 to +36 of L-DNA are identical to positions 1425_H to 1444_H and 1_H to 16_H of H-DNA, positions +37 to +82 correspond to positions 1129_H to 1174_H (one mismatch), and positions +83 to +116 are homologous to positions 101_H to 135_H (six mismatches and a 1-bp deletion). H-DNA base pair positions are given as numbers with a subscript H and refer to the standard H-DNA repeat. The termini of the linear genome are blunt ended with a 5'-phosphate group and are probably the result of site-specific cleavage between positions 962_H and 963_H (8, 101). The coding L-DNA contains three internal repeats; the leftmost repeat (19-bp [A] and 15-bp [B] motifs in ABABA configuration) is noncoding, while the other repeats (0.9 and 0.55 kb) are probably coding (Fig. 2).

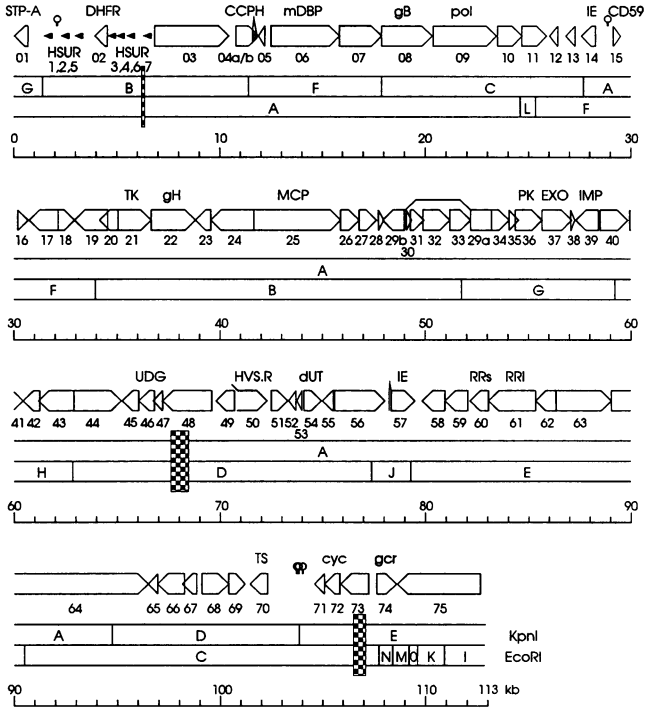

FIG. 2. Genetic map of herpesvirus saimiri L-DNA. Above the scale (in kilobases), a restriction map of the herpesvirus saimiri L-DNA for restriction endonucleases *Eco*RI and *Kpn*I is shown. Repeat sequences are indicated by pattern-filled rectangles within the map. Protein-coding regions are numbered from 01 to 75 and are displayed as open arrows. Protein-coding regions known to be encoded by spliced transcripts are connected with lines. In the case of ORF 50, it is assumed that a methionine codon is provided by a 5' exon, but the splice sites are not known. Genes for herpesvirus saimiri U-RNAs (HSUR) are shown as black arrowheads. Small circles upstream of genes 01, 14 and 15, and 70 indicate dyad symmetries. Abbreviations: CCPH, complement control protein homolog; mDBP, major DNA binding protein; pol, DNA polymerase; IE, immediate-early genes (ORFs 14 and 57); CD59, homolog of human CD59; TK, thymidine kinase; MCP, major capsid protein; PK, protein kinase; EXO, alkaline exonuclease; IMP, integral membrane protein; UDG, uracil DNA glycosylase; HVS.R, herpesvirus saimiri equivalent of EBV R *trans* activator; dUT, dUTP nucleotidohydrolase (dUTPase); RRs and RRl, small and large subunits, respectively, of the ribonucleotide reductase; cyc, cyclin homolog; gcr, G protein-coupled receptor homolog.

Nucleotide sequence accession number. The nucleotide sequence data from this study have been deposited in the EMBL, GenBank, and DDBJ nucleotide sequence data bases under accession number X64346.

TABLE 1. Nomenclature, position, and orientation of protein-coding regions of herpesvirus saimiri L-DNA

ORF[a]	Nucleotide position[b]		Strand[c]	No. of aa[d]	Mol mass[e]	ORF	Nucleotide position		Strand	No. of aa	Mol mass
	Start	End					Start	End			
01	223	714	C	164	17.1	38	57200	57397		66	7.5
02	3972	4532	C	187	21.7	39	57428	58525	C	366	42.2
03	6821	10558		1,246	138.9	40	58610	59959		450	51.7
04a	10912	11991		360	40.0	41	60095	60577		161	18.7
04b	10912	11773 (SD)		287	33.4	42	60575	61369	C	265	30.0
	11967 (SA)	12010		15		43	61338	63026	C	563	63.9
05 (KFLF1)	11998	12285	C	96	11.1	44	63001	65343		781	88.1
06 (KFRF1)	12584	15967		1,128	127.5	45	65387	66157	C	257	28.3
07 (KFRF2)	15977	18013		679	77.6	46	66171	66926	C	252	28.9
08 (KCRF1)	18006	20429		808	91.7	47	66895	67317	C	141	16.0
09 (KCRF2)	20507	23533		1,009	113.9	48 (EDLF5)	67326	69716	C	797	88.9
10 (KCRF3)	23573	24793		407	45.3	49 (EDLF4)	69928	70836	C	303	35.7
11 (KCRF4)	24799	26013		405	45.7	50 (EDRF1)	70798	72402		535	60.0
12 (KCLF1)	26048	26554	C	169	19.5	51 (EDRF2)	72626	73432		269	29.6
13 (KCLF2)	26939	27391	C	151	17.2	52 (EDLF3)	73471	73815	C	115	13.2
14	27677	28423	C	249	28.3	53 (EDLF2)	73860	74129	C	90	10.4
15	29231	29593		121	13.8	54 (EDRF3)	74206	75066		287	32.5
16	30322	30801		160	18.0	55 (EDLF1)	75107	75706	C	200	22.3
17	30840	32264	C	475	52.8	56 (EDRF4)	75685	78189		835	96.1
18	32257	33024		256	29.7	57a	78291	78309 (SD)		6	46.8
19	33027	34655	C	543	61.5	57b (EJRF1)	78396 (SA)	79624		410	
20	34288	35196	C	303	34.9	58 (EELF5)	79989	81059	C	357	40.5
21	35195	36775		527	59.8	59 (EELF4)	81059	82162	C	368	40.5
22	36775	38925		717	82.6	60 (EELF3)	82276	83190	C	305	35.2
23	38924	39682	C	253	28.9	61 (EELF2)	83199	85499	C	767	87.2
24	39695	41887	C	731	83.6	62 (EELF1)	85502	86491	C	330	37.4
25	41893	46005		1,371	154.4	63 (EERF1)	86498	89194		899	103.3
26	46024	46935		304	34.3	64 (EERF2)	89200	96606		2,469	280.2
27	46945	47784		280	32.7	65 (ECLF7)	96616	97032	C	139	15.3
28	47862	48140		93	10.4	66 (ECLF6)	97028	98332	C	435	50.1
29b	48092	49231 (SA)	C	380	77.0	67 (ECLF5)	98248	98952	C	235	26.9
30	49249	49473		75	8.3	68 (ECRF1)	99198	100505		436	49.1
31	49443	50066		208	24.3	69 (ECRF2)	100511	101293		261	29.8
32	50015	51337		441	50.8	70 (ECLF4)	101536	102417	C	294	33.5
33	51333	52322		330	37.0	71 (ECLF3)	104689	105189	C	167	19.1
29a	52315 (SD)	53223	C	303	77.0	72 (ECLF2)	105193	105954	C	254	28.6
34	53222	54169		316	36.3	73 (ECLF1)	106013	107233	C	407	46.6
35	54159	54608		150	13.3	74 (ECRF3)	107732	108694		321	37.1
36	54499	55791		431	48.9	75 (EILF1)	108785	112681	C	1,299	143.2
37	55794	57242		483	55.5						

[a] Protein-coding regions (ORFs) of herpesvirus saimiri are ordered by their position within the L-DNA and numbered from 01 to 75. Designations of genes described earlier are given in parentheses (2, 81, 82). Genes known to be spliced were supplied with the same number. In the case of gene 04, there exists an unspliced mRNA allocated to ORF 04a and a spliced transcript made of two exons which was assigned to ORF 04b (3). Gene 57 has been shown to contain two exons (82). Gene 29 has two exons, one each from ORF 29a and 29b; the splice junctions of this gene are conserved among sequenced mammalian herpesviruses (Fig. 3).

[b] The translation range (from ATG to the nonsense codon) is given as the position of an ORF. Exons of spliced mRNAs are separated and marked by SD (splice donor) and SA (splice acceptor).

[c] C indicates that the ORF is on the complementary strand.

[d] The number of amino acids (aa) is given for each ORF. In the case of ORF 50, the total number of codons is displayed because the first methionine codon occurs far into the ORF, although the similarity of the deduced amino acid sequence to the EBV BRLF1 gene product begins close to their amino termini.

[e] The calculated molecular mass was derived from the predicted proteins and is given in kilodaltons. In the case of ORF 50, the molecular mass was computed from the total number of codons.

RESULTS

Primary structure of L-DNA. Both junctions between H-DNA repeats and the termini of L-DNA have been defined previously (79, 101). The left H-DNA–L-DNA border is not clearly recognized, since a 116-bp G+C-rich DNA fragment with homology to repeat DNA lies between concatemeric H-DNA repeats and the A+T-rich L-DNA sequences. Using the numbering system of Murthy et al. (79), we define the first nucleotide that digresses from the standard H-DNA repeat unit as the beginning of L-DNA (position +1 [Fig. 1]). In contrast, the right L-to-H transition could be localized to a single nucleotide (position 962_H [Fig. 1]). It nearly coincides with an incomplete H-DNA packaging cleavage site (9, 101). On this basis, the entire L-DNA region of herpesvirus saimiri comprises 112,930 bp and has a G+C content of 34.5%.

Computer analysis predicted 76 open reading frames (ORFs) in the L-DNA of herpesvirus saimiri (Fig. 2). The criteria defining a potential ORF included (i) a minimum of 60 amino acids in the derived polypeptide, (ii) no more than a 60% overlap with other reading frames, (iii) the presence of typical translational start signals (59), (iv) potential promoter and transcriptional terminator elements, (v) codon preference similar to those of unambiguously identified viral genes,

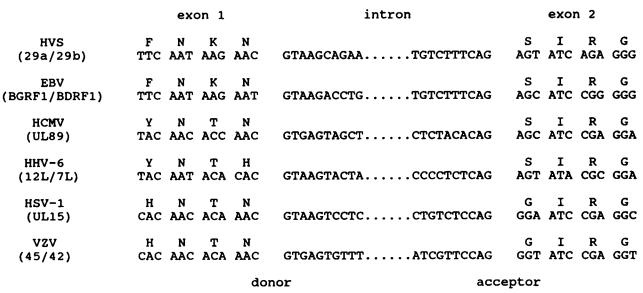

FIG. 3. Nucleotide and deduced amino acid sequences in the region of the predicted splice donor site of exon 1 and the predicted splice acceptor site of exon 2 of gene 29 and its counterparts in the human herpesviruses EBV (BGRF1/BDRF1) (6), HCMV (UL89) (18), human herpesvirus 6 (HHV-6) (12L/7L) (64), herpes simplex virus type 1 (HSV-1) (UL15) (23, 34, 69), and varicella-zoster virus (VZV) (gene 45/42) (27). Encoded protein sequences are shown in the single-letter code above the codons. The DNA sequences of the predicted splice sites conform to the GT-AG rule and to the consensus sequence for splice donor and acceptor sites (76). The splice junction of the HSV-1 UL15 gene has been verified experimentally (34).

and (vi) sequence homologies to known reading frames of other herpesviruses and cellular genes. Table 1 lists the ORFs that are likely to encode proteins, including the number of codons and the calculated molecular masses of the predicted proteins. The well-conserved reading frames 29a and 29b are listed separately. Since, however, the homologous ORFs of herpes simplex virus were shown to be spliced (23, 34), it is assumed that analogous processing of gene 29 results in a 77-kDa herpesvirus saimiri protein (Fig. 3). ORF 04 has been shown to be transcribed into two variants. An unspliced mRNA encodes a viral membrane glycoprotein, while a spliced mRNA directs translation of a secreted glycoprotein with the structure of a complement-regulating protein (3). Thus, in the absence of more data on transcription and splicing, we can predict 76 distinct herpesvirus saimiri-encoded proteins. Usually, the reading frames are closely arranged, and they frequently overlap. Sizable stretches (>600 bp) of noncoding L-DNA are found only in the promoter and 5'-untranslated region of the TS gene (ORF 70); upstream of ORF 15; upstream of ORF 14, which is expressed under immediate-early conditions (85); and in the left-terminal L-DNA, which encompasses seven U-RNA genes termed herpesvirus saimiri U-RNAs (HSURs) (4, 78, 110) (two such stretches are found in the left-terminal L-DNA). Three of these regions bear dyad symmetries similar to those found in the origins of DNA replication of other herpesviruses (40, 103, 116). Three regions have characteristic clusters of short tandem repetitions. Five repeat elements (19 bp [A] and 15 bp [B] in an ABABA configuration) are located in the noncoding region downstream of HSUR 7; two repeat clusters of 0.9 and 0.55 kb are located within ORFs 48 and 73, respectively (Fig. 2).

Collinear genome organization of herpesvirus saimiri and EBV, a subgroup 1 gammaherpesvirus. Despite differences in the overall genome structure of herpesvirus saimiri and EBV, the majority of herpesvirus saimiri genes have counterparts in the EBV genome at the level of amino acid sequence similarities (Table 2), and all of these genes are arranged in collinear order (Fig. 4). We note that herpesvirus saimiri ORFs 18, 30, and 35 are homologous to sequences in the EBV DNA that had not been listed as separate reading frames by Baer et al. (6). Herpesvirus saimiri ORF 11 is homologous to ORF Raji LF2, which occurs in the EBV episome of Raji cells but is deleted in B95-8 genomes (89). The differences between EBV and herpesvirus saimiri include the known genes relevant for transformation and persistence, including all EBNA genes of EBV, the coding regions for the latent membrane protein (LMP), and the terminal protein (6, 63, 98). Likewise, the STP and HSUR genes located in the left-terminal L-DNA of herpesvirus saimiri do not have counterparts in the EBV genome (Fig. 4). ORFs 02 (DHFR) (108), 14 (immediate-early G [IE-G]) (83, 85), 15 (homolog of human CD59) (5), 70 (TS) (14, 45), and 72 (ECLF2, a cyclin homolog) (82), among others, appeared unique to herpesvirus saimiri, while the well-characterized EBV genes for glycoproteins gp350 and gp220 and EBV-encoded small RNAs (EBERs) (96) are not found in herpesvirus saimiri.

Amino acid sequence homology with alphaherpesvirus and betaherpesviruses. Sequence analysis of herpesvirus saimiri has detected amino acid sequences homologous to those of the four human herpesviruses which have been completely sequenced (6, 18, 27, 69, 70, 89). Table 3 shows that human herpesvirus genes can be grouped into at least six blocks that have remained intact and mostly collinear with respect to encoded proteins throughout the evolution of the known mammalian herpesvirus subgroups. Block I, which contains ORFs 06 to 09 (encoding the major DNA binding protein, a putative transport protein, the glycoprotein B [gB] equivalent, and the DNA polymerase, respectively), represents the longest stretch of contiguous genes with pronounced sequence homologies. Block II, which includes, among other genes, those encoding thymidine kinase and gH, is not always intact; for instance, HCMV lacks a thymidine kinase. Gene block III (which includes the major capsid protein gene) is always located between blocks II and IV; however, its orientation in the betaherpesviruses HCMV and human herpesvirus 6 is opposite to the direction of the corresponding alphaherpesvirus and gammaherpesvirus virus block (18, 64). Block V, which has genes of limited homology, includes the dUTPase gene and ORF 57, which encodes the immediate-early transactivator gene product (84), which has similarity in structure and function to the EBV transactivator encoded by BMLF1 (6) and the immediate-early protein encoded by gene UL54 of herpes simplex virus (69). Block VI contains the reading frames for the large and small subunits of the ribonucleotide reductase, which are highly conserved between alphaherpesviruses and gammaherpesviruses; however, this enzyme is only partially conserved in the betaherpesvirus HCMV, which does not encode the catalytic small subunit of the ribonucleotide reductase (18).

Homologs of cellular genes in herpesvirus saimiri. Sequence analysis of the herpesvirus saimiri L-DNA by searches of current versions of sequence data bases allowed the detection of more genes with striking structural similarities to cellular genes than for any other known virus. Besides two genes for enzymes of nucleotide metabolism, TS (14, 45) and DHFR (11, 108), we have identified two ORFs for proteins that, on the basis of structural characteristics, should be down-regulators of the complement system. Herpesvirus saimiri ORF 04 can be transcribed into an unspliced mRNA that codes for a 65- to 75-kDa virion surface glycoprotein (3) with about 30% amino acid sequence identity to most human membrane-bound complement control proteins, including membrane cofactor protein (CD46) (67), decay-accelerating factor (CD55) (17, 72), and complement receptor types 1 (CD35) (54) and 2 (CD21) (111). Splicing creates a further

TABLE 2. Detectable homologies of herpesvirus saimiri translation products with EBV protein sequences[a]

HVS[b] ORF	No. of codons	EBV ORF[c]	No. of codons	% Identity[d]	% Similarity[d]
03	1,246	BNRF1[e]	1,318	27.6	49.8
06	1,128	BALF2	1,128	41.4	63.5
07	679	BALF3	789	39.1	59.9
08	808	BALF4	857	41.1	60.0
09	1,009	BALF5	1,015	54.8	70.3
11	405	Raji LF2[f]	429	30.7	50.2
17	475	BVRF2	605	38.0	57.4
18[g]	256	BVLF1.5a/b	246	30.6	55.5
19	543	BVRF1	570	37.9	58.2
20	303	BXRF1	248	34.0	49.0
21	527	BXLF1	607	30.3	49.7
22	717	BXLF2	706	26.8	51.0
23	253	BTRF1	425	27.3	48.6
24	731	BcRF1	618	39.2	59.3
25	1,371	BcLF1	1,381	57.6	74.6
26	304	BDLF1	302	48.0	71.7
27[h]	280	BDLF2	420	18.1	44.5
29b	380	BDRF1	387	57.0	75.6
30[g]	75	BDLF3.5	77	28.1	54.7
31	208	BDLF4	225	38.7	61.8
32	441	BGLF1	507	28.6	47.9
33	330	BGLF2	336	42.5	60.6
29a	303	BGRF1	325	44.9	63.7
34	316	BGLF3	332	34.6	58.7
35[g]	150	BGLF3.5	153	24.6	50.7
36	431	BGLF4	455	31.0	51.5
37	483	BGLF5	470	42.0	62.3
38	66	BBLF1	75	28.8	53.0
39	366	BBRF3	405	50.0	72.2
40	450	BBLF2	541	21.9	47.6
42	265	BBRF2	278	35.4	57.3
43	563	BBRF1	613	52.0	68.0
44	781	BBLF4	809	52.7	68.3
45	257	BKRF4	214	30.0	45.7
46	252	BKRF3	255	52.8	70.4
47	141	BKRF2	137	25.6	49.6
48[h]	797	BRRF2	537	18.0	43.2
49	303	BRRF1	310	22.3	43.2
50	535	BRLF1	605	22.7	43.5
52	115	BLRF2	162	28.7	53.9
53	90	BLRF1	102	27.8	52.2
54	287	BLLF3	278	28.3	52.2
55	200	BSRF1	218	46.3	66.2
56	835	BSLF1	874	37.5	55.9
57	416	BMLF1	438	27.2	46.1
58	357	BMRF2	357	25.5	49.9
59	368	BMRF1	404	33.1	54.7
60	305	BARF1	302	60.3	73.5
61	767	BORF2	826	42.8	61.1
62	330	BORF1	364	30.4	57.1
63	899	BOLF1	1,239	25.5	44.2
64	2,469	BPLF1	3,149	24.6	45.8
65	139	BFRF3	176	30.4	51.4
66	435	BFRF2	591	29.0	49.6
67	235	BFRF1	336	41.6	61.4
68	436	BFLF1	525	38.1	62.4
69	261	BFLF2	318	39.2	60.0
75	1,299	BNRF1[e]	1,318	28.8	51.2

[a] Horizontal rules indicate areas of contiguous collinearity which are separated by nonhomologous genes (see also Fig. 4).
[b] HVS, herpesvirus saimiri.
[c] EBV ORF nomenclature follows the conventions described by Baer et al. (6).
[d] Amino acid sequences were compared by using the GAP option of the Genetics Computer Group software package with the gap weight and gap length parameters usually set to 3.0 and 0.1, respectively. Amino acid sequence identity and similarity, based on conservative substitutions (28), are given for homologous protein sequences.
[e] The EBV gene product of BNRF 1 is the only one that has two homologs in herpesvirus saimiri (ORFs 03 and 75).
[f] The Raji LF2 ORF was found only in the EBV episome of Raji cells but is deleted in the EBV prototype genome B95-8 (89).
[g] Comparisons involving herpesvirus saimiri ORFs 18, 30, and 35 identified homologous protein sequences in EBV that have not been reported previously. In the case of herpesvirus saimiri ORF 18, two overlapping ORFs of EBV (BVLF 1.5a and BVLF 1.5b, indicated here as BVLF 1.5 a/b) displayed significant homologies; the number of codons for BVLF 1.5a/b resulted from a hypothesized frameshift elimination.
[h] ORFs 27 and 48 of herpesvirus saimiri have limited (<20% identity) homology to the corresponding EBV ORFs; the significance of these findings is uncertain.

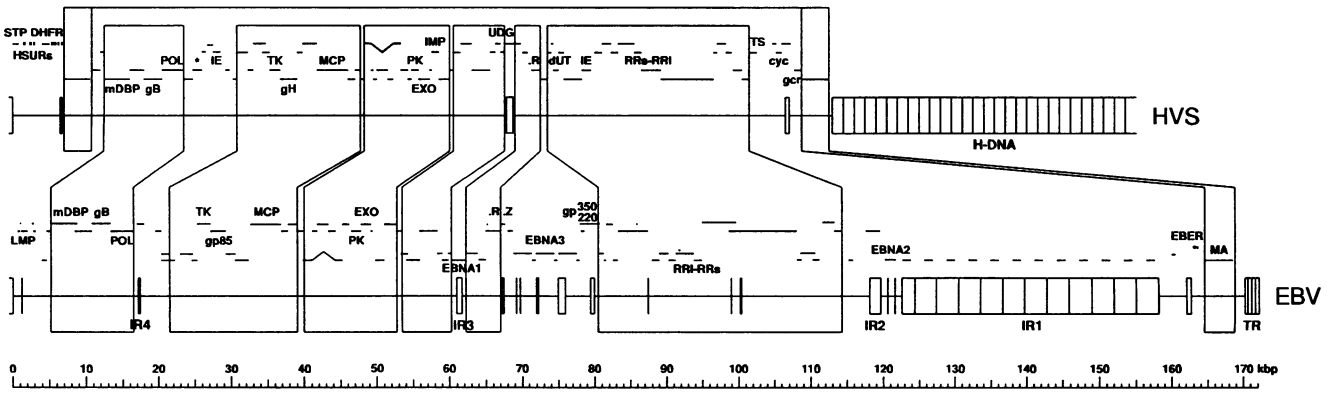

FIG. 4. Collinearity of herpesvirus saimiri (HVS) and EBV. The genomes of HVS and EBV are displayed as horizontal lines interspersed with open boxes indicating repetitive DNA sequences. The genome of EBV has been inverted relative to its conventional orientation (6). Above the genomes, ORFs are symbolized by short lines and dots. Seven blocks of genes conserved among HVS and EBV are boxed. Abbreviations are as in the legend to Fig. 2; additional abbreviations are as follows: LMP, latent membrane protein; IR, internal repeat; EBNA, Epstein-Barr virus nuclear antigen; R, R *trans* activator; Z, Z *trans* activator; MA, membrane antigen (16); TR, terminal repeat.

transcript of gene 04 that specifies a secreted glycoprotein of 47 to 53 kDa (3) related to the soluble complement inhibitors C4b-binding protein (21) and factor H (93) (Table 4). In addition, ORF 15 resembles another complement regulator gene; it is 49% identical to the gene for human surface glycoprotein CD59 (also known as membrane attack complex inhibitory factor [MACIF] or the 20-kDa homologous restriction factor [HRF20]) (5, 26, 87, 100, 104). The transformation-associated gene product STP-A (STP of subgroup A) has a number of collagenlike repeat elements (Gly-X-Y, where X, Y, or both are prolines); this is even more apparent in the STP of herpesvirus saimiri subgroup C strain 488, where the repeat elements have been shown to be contiguous (13, 51). Three other ORFs could encode proteins related to proliferation control. ORF 14 represents one of two characterized immediate-early genes specifying *trans*-acting functions (83–85). The protein is related in sequence to the Mls genes in mice (Table 4) (85, 107), which are superantigen gene prototypes (1, 20). As described recently (82), the predicted polypeptide encoded by ORF 72 (ECLF2) is homologous to members of the cyclin family of proteins, being most closely related to the D-type cyclins which are involved in the regulation of the G_1-to-S phase transition of eucaryotic cells (68, 114, 115). The translation product of ORF 74 (ECRF3) (84) of herpesvirus saimiri resembles in structure the products of ORFs for G protein-coupled receptors, e.g., it shows 30% identity with the human low affinity interleukin-8 receptor, which is a powerful activating factor for neutrophils (77). The ORF 74 gene product is also homologous to the US27-, US28-, and UL33-specified proteins of HCMV (18, 19), but the more closely related gammaherpesvirus EBV does not contain a counterpart of this gene.

A number of polypeptides specified by herpesvirus saimiri genes contain structural motifs that are typically conserved in functionally related enzymes and protein families (Table 5). While most of these are shared by all known herpesviruses, the product of ORF 27 displays, at most, limited amino acid sequence similarity to the EBV BDLF2-encoded protein but no similarity to other herpesvirus proteins. The ORF 27-derived gene product was recognized by an amino acid signature identifying cytosine-specific methylases (10, 62). This may be relevant to the observation that persisting episomal herpesvirus saimiri DNA in tumor cell lines (31) and lymphoid cells of tumor-bearing animals (29) is highly methylated in a characteristic and well-conserved pattern. The amino acid sequence deduced from ORF 36 was targeted by a search for protein kinase motifs (7). If the two sequence motifs shown in Table 5 are included in one protein, the probability that this protein is a protein kinase is close to 100%. The larger of the two motifs is consistently found in tyrosine-specific protein kinases (41). The equivalent protein sequence of EBV has this specificity, while the related polypeptides of alphaherpesviruses and betaherpesviruses appear to be serine/threonine-specific phosphorylating enzymes.

DISCUSSION

Completion of the determination of the herpesvirus saimiri nucleotide sequence revealed a number of novel observations. Herpesvirus saimiri L-DNA has a total of 112,930 bp, including a short G+C-rich transition at the left end, and is flanked by approximately 35 G+C-rich (70.8%) tandem repeats of 1,444 bp each. The G+C content of L-DNA is 34.5%, and it is markedly depleted in CpG, while this dinucleotide occurs mostly at random frequencies in alphaherpesviruses and betaherpesviruses. This may be related to methylation of episomal herpesvirus saimiri DNA persisting in T cells or their derivatives (2, 31, 47). Similarly, EBV DNA is reduced in CpG content, and it is found methylated in some, but not all, lymphoblast cell lines in which EBV persists (53, 61). 5-Methylcytosine is a hot spot for mutations, as deamination of 5-methylcytosine causes a change to thymidine (24). The dinucleotide CpG might have been lost specifically, since methylation activity is particularly high in lymphoid cells containing herpesvirus saimiri (29, 31). We have now found that ORF 27 has the characteristic signature of a cytosine-specific methylase. Experiments are planned to determine its expression in persistently infected cells and to verify the enzymatic activity predicted from the 5-methylcytosine–methylase motif. In addition, herpesvirus saimiri encodes two enzymes involved in dTMP synthesis (TS and DHFR) (11) that may contribute to the high A+T content of L-DNA (45), although this feature may also be a consequence of DNA methylation and reactivation of persisting genomes. Herpesvirus saimiri is less genetically complex than the other fully characterized herpesviruses, and there

TABLE 3. Conserved genes of herpesvirus saimiri and human herpesviruses[a]

HVS[b] ORF	Corresponding gene in[c]:				Comment(s)[d]	Reference(s)
	EBV	HCMV	VZV	HSV-1		
03	BNRF1				Homologous to HVS ORF 75	
06	BALF2	UL57	29	UL29	mDBP	2
07	BALF3	UL56	30	UL28	Transport protein	2, 90
08	BALF4	UL55	31	UL27	gB	2
09	BALF5	UL54	28	UL30	DNA polymerase	2
11	Raji LF2					89
17	BVRF2	UL80	33	UL26	Proteinase, assemblin	112, 113
18	BVRF1.5 a/b	UL79				
19	BVRF1	UL77	34	UL25	Virion protein	69
20	BXRF1	UL76	35	UL24		
21	BXLF1		36	UL23	TK	46
22	BXLF2	UL75	37	UL22	gH	38
23	BTRF1					
24	BcRF1	UL87				
25	BcLF1	UL86	40	UL19	MCP	18, 27, 69
26	BDLF1	UL85	41	UL18		
27	BDLF2					
29b	BDRF1	UL89	42	UL15	Spliced ORF, exon 2	23, 34
30	BDLF3.5	UL91				
31	BDLF4	UL92				
32	BGLF1	UL93	43	UL17		
33	BGLF2	UL94	44	UL16		
29a	BGRF1	UL89	45	UL15	Spliced ORF, exon 1	23, 34
34	BGLF3	UL95	46 (p)	UL14 (p)		
35	BGLF3.5	UL96				
36	BGLF4	UL97	47	UL13	Putative protein kinase	
37	BGLF5	UL98	48	UL12	Alkaline exonuclease	7
38	BBLF1	UL99 (p)	49	UL11		
39	BBRF3	UL100	50	UL10	Integral membrane protein	66
40	BBLF2	UL101 (p)	51 (p)	UL9 (p)		
41	BBLF3 (p)	UL102 (p)	52 (p)	UL8 (p)		
42	BBRF2	UL103	53	UL7		
43	BBRF1	UL104	54	UL6	Virion protein	69
44	BBLF4	UL105	55	UL5	Helicase	42, 43, 117
45	BKRF4					
46	BKRF3	UL114	59	UL2	Uracil DNA glycosylase	88
47	BKRF2					
48	BRRF2					
49	BRRF1					
50	BRLF1				HVS.R	83
52	BLRF2					
53	BLRF1	UL73				
54	BLLF3	UL72 (p)	8	UL50	dUTPase	91
55	BSRF1	UL71 (p)	7	UL51 (p)		
56	BSLF1	UL70	6	UL52	Helicase-primase complex	25
57	BMLF1	UL69	4	UL54	IE-52k	84
58	BMRF2		15 (p)	UL43 (p)		
59	BMRF1		16 (p)	UL42 (p)		
60	BARF1		18	UL40	RR small subunit	86
61	BORF2	UL45	19	UL39	RR large subunit	86
62	BORF1	UL46	20	UL38	Virion protein	69
63	BOLF1	UL47	21 (p)	UL37 (p)		
64	BPLF1	UL48	22	UL36	Virion protein	69
65	BFRF3		23 (p)	UL35 (p)		
66	BFRF2	UL49				
67	BFRF1	UL50	24 (p)	UL34 (p)		
68	BFLF1	UL52	26	UL32		
69	BFLF2	UL53	27	UL31		
70			13		TS	14, 45, 106
74		US28			G protein-coupled receptor	19, 82
75	BNRF1				160-kDa virion protein	16

[a] Boxed areas indicate six blocks of genes are conserved among sequenced human herpesviruses and herpesvirus saimiri.

[b] HVS, herpesvirus saimiri.

[c] Homologies were identified by using the FASTA or GAP option of the Genetics Computer Group program package (33). p indicates viral genes that are positionally, and in some cases functionally, analogous to those of herpesvirus saimiri but have very limited, if any, sequence similarity. The nomenclature of gene designations was taken from descriptions of complete genomic sequences (6, 18, 27, 69). VZV, varicella-zoster virus; HSV-1, herpes simplex virus type 1.

[d] mDBP, major DNA binding protein; TK, thymidine kinase; MCP, major capsid protein; HVS.R, herpesvirus saimiri equivalent of EBV R *trans* activator; RR, ribonucleotide reductase.

TABLE 4. Sequence homology of herpesvirus saimiri-encoded proteins and cellular proteins

ORF	Designation	Comment(s)[a]	Amino acid sequence identity	Reference(s)
01	STP-A	Functionally related to STP-C (strain 488)	Local only, collagenlike repeats	51
02	DHFR		83% (human DHFR)	108
04 a/b	CCPH[b]	Striking similarities with genes for:		3
		Membrane cofactor protein (CD46)	29%	67
		Decay-accelerating factor (CD55)	28%	17, 72
		C4b-binding protein	34%	21
		Complement receptor type 1(CD35)	27%	54
		Complement receptor type 2 (CD21)	30%	111
		Factor H	26%	93
		Major secretory protein of vaccinia virus	37%	57, 58
14	IE-G	Possibly a superantigen gene		107
		Homologous to Mls genes in mice and MMTV LTR[c] 3′ ORF	<43%, local only	1, 20, 85
15		Homologous to human CD59	49%	5, 87, 99, 104
70	TS		66% (human TS)	14, 45
72	ECLF2	Homologous to members of the cyclin family	25% (cyclin D)	82, 115
74	ECRF3	Homologous to G protein-coupled receptors	22% (HCMV US28)	19, 82
			30% (human IL-8[d] receptor)	77

[a] A TS has also been found in varicella-zoster virus (106), and homologs of G protein-coupled receptors have been identified in HCMV (19). None of the other proteins indicated has been identified in any other herpesvirus.
[b] CCPH, complement control protein homolog.
[c] MMTV LTR, mouse mammary tumor virus long terminal repeat.
[d] IL-8, interleukin 8.

are only very limited intergenic regions in L-DNA. The longest stretches devoid of reading frames are at the left end of L-DNA, where herpesvirus saimiri has seven HSUR genes, including promoters with typical transcription regulator elements of cellular U-RNAs (4, 13, 65, 78, 110). The HSUR genes are not expressed in productively infected cells but are expressed in transformed cells. It may be hypothesized that the HSURs contribute to cell transformation by attenuating the degradation of cellular mRNAs involved in T-lymphocyte growth regulation (80). Besides HSURs, at least four other viral genes are candidates as contributing factors for stimulation of T-cell proliferation. ORF 01 (STP-A) is necessary for in vitro immortalization (30, 56, 79) and has the ability to transform cells of the established rodent line Rat-1 (51). ORF 14 (IE-G) is similar to an ORF in the long terminal repeat of mouse mammary tumor virus and Mls genes of mice (85, 107). ORF 72 (ECLF2) (82) is homologous to members of the cyclin family and is most closely related to human cyclin D1 (115), which is the PRAD1 candidate oncogene and may play an important role

TABLE 5. Amino acid sequence motifs conserved between herpesvirus saimiri proteins and functionally characterized proteins

ORF	Motif (from PROSITE data base)[a]	Herpesvirus saimiri motif[a]	Identification[b]	Reference(s)
02	(L,I,F)Gx4(L,I,V,M,F)PW	IGKQGNLPW	DHFR	108
09	(Y,A)xDTDS(L,I,V,M,T)	YGDTDSL	DNA polymerase	2
21	(A,G)x4GK(S,T)	GSIGVGKT	ATP-A(TK)	46, 109
27	(R,K,Q)x2GN(S,T,A)(L,I,V,M)x3 (L,I,V,M)x3(L,I,V,M)x3(L,I,V,M)	RSPGNSVLGGLGQRVPRTV	^{5}mC methylase	10, 62
32	(A,G)x4GK(S,T)	GVHTLGKS	ATP-A (?)	109
36	(L,I,V)GxGx(F,Y)(S,G)x(L,I,V)	LGSGSFGSV	Protein kinase ATP	52
	(L,I,V,M,F,Y,C)x(H,Y)xD (L,I,V,M,F,Y)(R,S,T,A)x 2N(L,I,V,M,F,C)3	IIHSDISTSNILV	Protein kinase (Tyr)	41
44	(A,G)x4GK(S,T)	GTAGAGKS	ATP-A (helicase)	42, 43, 109, 117
46	WAx2GVL3N	WATQGVLLLN	Uracil DNA glycosylase	88
54	(A,G)x4GK(S,T)	AFILYGKS	ATP-A (dUTPase)	91, 109
60	Ex(L,I,V)Hx3Yx2(L,I,V)x3 (L,I,V,M,F,Y)3	ENIHGKVYANILNMLF	Ribonucleotide reductase small subunit	86
61	Gx2NSx3AxMP	GVFNSQFIALMP	Ribonucleotide reductase large subunit	86
70	LxPC(H,A,V)x3(Q,M)(F,Y,W)xV	LPPCHVLSQFYV	TS	14, 45
72	Rx2(L,I,V,M)x2(F,Y,W)(L,I,V,M)x8(L,I,V, M)x4(L,I,V,M,F,Y)x2(S,T,A,G)(L,I,V, M,F,Y)x(L,I,V,M,F,Y)2D(R,K)(L,I,V,M, F,Y,W)	RTILLTWMHLLCESFELDK SVFPLSVSILDRY	Cyclin	82

[a] Amino acid symbols are in the standard one-letter code, amino acids in parentheses are alternatives, x is any amino acid, and a number after a residue indicates its quantity.
[b] The ATP-A motif is characteristic for a number of ATP- and GTP-binding proteins and is generally referred to as the A consensus sequence (109). In the case of ORF 32, the relevance of the identified ATP-A motif is unclear; otherwise, the proposed function of a protein with an ATP-A motif is given in parentheses. TK, thymidine kinase; ^{5}mC, 5-methylcytosine.

in the development of centrocytic lymphoma and certain parathyroid tumors (75, 97). ORF 74 (ECRF3) (82) has the structure of G protein-coupled receptor molecules which have been shown to be involved in malignant transformation (48, 49). Also worth noting is the presence of two genes (ORF 04 and ORF 15) (3, 5) encoding glycoproteins related to complement control proteins which down-regulate complement activity at two distinct steps of complement activation (22, 60, 92). This suggests that herpesvirus saimiri may escape from host immune defenses by virtue of these viral gene products and thereby prevent elimination of a virus that is capable of persisting in peripheral blood lymphocytes. With at least 15 genes with pronounced homology to cellular DNA or proteins, the genome of herpesvirus saimiri, and possibly those of other subgroup 2 gammaherpesviruses, seems to be particularly prone to sequestering cellular genes. The genome appears to function as a spontaneous vector for cellular genes that may have been acquired by a mechanism involving reverse transcription, since most of the viral counterparts have no introns. The uptake of such genes may provide functions necessary for the progression of biological properties and secure a selection advantage in the natural host.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft (Forschergruppe "DNA-Viren des hämatopoetischen Systems").

REFERENCES

1. **Acha-Orbea, H., and E. Palmer.** 1991. Mls—a retrovirus exploits the immune system. Immunol. Today **12**:356–361.
2. **Albrecht, J.-C., and B. Fleckenstein.** 1990. Structural organization of the conserved gene block of herpesvirus saimiri coding for DNA polymerase, glycoprotein B, and major DNA binding protein. Virology **174**:533–542.
3. **Albrecht, J.-C., and B. Fleckenstein.** 1992. New member of the multigene family of complement control proteins in herpesvirus saimiri. J. Virol. **66**:3937–3940.
4. **Albrecht, J.-C., and B. Fleckenstein.** 1992. Nucleotide sequence of HSUR 6 and HSUR 7, two small RNAs of herpesvirus saimiri. Nucleic Acids Res. **20**:1810.
5. **Albrecht, J.-C., J. Nicholas, K. R. Cameron, C. Newman, B. Fleckenstein, and R. W. Honess.** Herpesvirus saimiri has a gene specifying a homolog of the cellular membrane glycoprotein CD59. Virology, in press.
6. **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, P. S. Tuffnell, and B. G. Barrell.** 1984. DNA sequence and expression of the B95-8 Epstein-Barr virus genome. Nature (London) **310**:207–211.
7. **Bairoch, A., and J.-M. Claverie.** 1988. Sequence patterns in protein kinases. Nature (London) **331**:22.
8. **Bankier, A. T., and B. G. Barrell.** 1983. Shotgun DNA sequencing, p. 1–33. *In* R. A. Flavell (ed.), Techniques in the life sciences, vol. B5. Elsevier, Amsterdam.
9. **Bankier, A. T., W. Dietrich, R. Baer, B. G. Barrell, F. Colbère-Garapin, B. Fleckenstein, and W. Bodemer.** 1985. Terminal repetitive sequences in herpesvirus saimiri virion DNA. J. Virol. **55**:133–139.
10. **Bestor, T., A. Laudano, R. Mattaliano, and V. Ingram.** 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzyme is related to bacterial restriction methyltransferases. J. Mol. Biol. **203**:971–983.
11. **Biesinger, B., R. Grassmann, B. Fleckenstein, S. C. S. Murthy, J. Trimble, and R. C. Desrosiers.** 1989. Genes for the synthesis of deoxythymidylate monophosphate in T-cell lymphoma-inducing herpesviruses of nonhuman primates. Curr. Top. Microbiol. Immunol. **144**:241–247.
12. **Biesinger, B., I. Müller-Fleckenstein, B. Simmer, G. Lang, S. Wittmann, E. Platzer, R. C. Desrosiers, and B. Fleckenstein.** 1992. Stable growth transformation of human T lymphocytes by herpesvirus saimiri. Proc. Natl. Acad. Sci. USA **89**:3116–3119.
13. **Biesinger, B., J. J. Trimble, R. C. Desrosiers, and B. Fleckenstein.** 1990. The divergence between two oncogenic herpesvirus saimiri strains in a genomic region related to the transforming phenotype. Virology **176**:505–514.
14. **Bodemer, W., H. H. Niller, N. Nitsche, B. Scholz, and B. Fleckenstein.** 1986. Organization of the thymidylate synthase gene of herpesvirus saimiri. J. Virol. **60**:114–123.
15. **Bridgen, A., A. J. Herring, N. F. Inglis, and H. W. Reid.** 1989. Preliminary characterization of the alcelaphine herpesvirus I genome. J. Gen. Virol. **70**:1141–1150.
16. **Cameron, K. R., T. Stamminger, M. Craxton, W. Bodemer, R. W. Honess, and B. Fleckenstein.** 1987. The 160,000-M_r virion protein encoded at the right end of the herpesvirus saimiri genome is homologous to the 140,000-M_r membrane antigen encoded at the left end of the Epstein-Barr virus genome. J. Virol. **61**:2063–2070.
17. **Caras, I. W., M. A. Davitz, L. Rhee, G. Weddell, D. W. Martin, Jr., and V. Nussenzweig.** 1987. Cloning of decay-accelerating factor suggests novel use of splicing to generate two proteins. Nature (London) **325**:545–549.
18. **Chee, M. S., A. T. Bankier, S. Beck, R. Bohni, C. M. Brown, R. Cerny, T. Horsnell, C. A. Hutchison III, T. Kouzarides, J. A. Martignetti, E. Preddie, S. C. Satchwell, P. Tomlinson, K. M. Weston, and B. G. Barrell.** 1990. Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. Curr. Top. Microbiol. Immunol. **154**:125–169.
19. **Chee, M. S., S. C. Satchwell, E. Preddie, K. M. Weston, and B. G. Barrell.** 1990. Human cytomegalovirus encodes three G protein-coupled receptor homologues. Nature (London) **344**:774–777.
20. **Choi, Y., J. W. Kappler, and P. Marrack.** 1991. A superantigen encoded in the open reading frame of the 3′ long terminal repeat of mouse mammary tumor virus. Nature (London) **350**:203–207.
21. **Chung, L. P., D. R. Bentley, and K. B. M. Reid.** 1985. Molecular cloning and characterization of the cDNA coding for the C4b-binding protein, a regulatory protein of the classical pathway of the human complement system. Biochem. J. **230**:133–141.
22. **Cooper, N. R.** 1991. Complement evasion strategies of microorganisms. Immunol. Today **12**:327–331.
23. **Costa, R. H., K. G. Draper, T. J. Kelly, and E. K. Wagner.** 1985. An unusual spliced herpes simplex virus type 1 transcript with sequence homology to Epstein-Barr virus DNA. J. Virol. **54**:317–328.
24. **Coulondre, C., J. H. Miller, P. J. Farabaugh, and W. Gilbert.** 1978. Molecular basis of base substitution hotspots in Escherichia coli. Nature (London) **274**:775–780.
25. **Crute, J. J., T. Tsurumi, L. Zhu, S. K. Weller, P. D. Olivo, M. D. Challberg, E. S. Mocarski, and I. R. Lehman.** 1989. Herpes simplex virus 1 helicase-primase: a complex of three herpes-encoded gene products. Proc. Natl. Acad. Sci. USA **86**:2186–2189.
26. **Davies, A., D. L. Simmons, G. Hale, R. A. Harrison, H. Tighe, P. J. Lachmann, and H. Waldmann.** 1989. CD59, an LY-6-like protein expressed in human lymphoid cells, regulates the action of the complement membrane attack complex on homologous cells. J. Exp. Med. **170**:637–654.
27. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella-zoster virus. J. Gen. Virol. **67**:1759–1816.
28. **Dayhoff, M. O., W. C. Barker, and L. T. Hunt.** 1983. Establishing homologies in protein sequences. Methods Enzymol. **91**:524–545.
29. **Desrosiers, R. C.** 1982. Specifically unmethylated cytidylic-guanylate sites in *Herpesvirus saimiri* DNA in tumor cells. J. Virol. **43**:427–435.
30. **Desrosiers, R. C., A. Bakker, J. Kamine, L. A. Falk, R. D. Hunt, and N. W. King.** 1985. A region of the herpesvirus

saimiri genome required for oncogenicity. Science **228:**184–187.
31. **Desrosiers, R. C., C. Mulder, and B. Fleckenstein.** 1979. Methylation of herpesvirus saimiri DNA in lymphoid tumor cell lines. Proc. Natl. Acad. Sci. USA **76:**3839–3843.
32. **Desrosiers, R. C., D. P. Silva, L. M. Waldron, and N. L. Letvin.** 1986. Nononcogenic deletion mutants of herpesvirus saimiri are defective for in vitro immortalization. J. Virol. **57:**701–705.
33. **Devereux, J., P. Haeberli, and O. Smithies.** 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. **12:**387–395.
34. **Dolan, A., M. Arbuckle, and D. J. McGeoch.** 1991. Sequence analysis of the splice junction in the transcript of herpes simplex virus type 1 gene UL15. Virus Res. **20:**97–104.
35. **Efstathiou, S., Y. M. Ho, and A. C. Minson.** 1990. Cloning and molecular characterization of the murine herpesvirus 68 genome. J. Gen. Virol. **71:**1355–1364.
36. **Falk, L. A., L. G. Wolfe, and F. Deinhardt.** 1972. Isolation of herpesvirus saimiri from blood of squirrel monkeys (Saimiri sciureus). J. Natl. Cancer Inst. **48:**1499–1505.
37. **Fleckenstein, B., and R. C. Desrosiers.** 1982. Herpesvirus saimiri and herpesvirus ateles, p. 253–332. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Press, New York.
38. **Gompels, U. A., M. A. Craxton, and R. W. Honess.** 1988. Conservation of glycoprotein H (gH) in herpesviruses: nucleotide sequence of the gH gene from herpesvirus saimiri. J. Gen. Virol. **69:**2819–2829.
39. **Gompels, U. A., M. A. Craxton, and R. W. Honess.** 1988. Conservation of gene organization in the lymphotropic herpesviruses herpesvirus saimiri and Epstein-Barr virus. J. Virol. **62:**757–767.
40. **Hamzeh, F. M., P. S. Lietman, W. Gibson, and G. S. Hayward.** 1990. Identification of the lytic origin of DNA replication in human cytomegalovirus by a novel approach utilizing ganciclovir-induced chain termination. J. Virol. **64:**6184–6195.
41. **Hanks, S. K., A. M. Quinn, and T. Hunter.** 1988. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. Science **241:**42–52.
42. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. Nature (London) **333:**22–23.
43. **Hodgman, T. C.** 1988. A new superfamily of replicative proteins. Nature (London) **333:**578. (Erratum.)
44. **Honess, R. W.** 1984. Herpes simplex and the 'the herpes complex': diverse observations and a unifying hypothesis. J. Gen. Virol. **65:**2077–2107.
45. **Honess, R. W., W. Bodemer, K. R. Cameron, H.-H. Niller, B. Fleckenstein, and R. E. Randall.** 1986. The A+T-rich genome of herpesvirus saimiri contains a highly conserved gene for thymidylate synthase. Proc. Natl. Acad. Sci. USA **83:**3604–3608.
46. **Honess, R. W., M. A. Craxton, L. Williams, and U. A. Gompels.** 1989. A comparative analysis of the sequence of the thymidine kinase gene of a gammaherpesvirus, herpesvirus saimiri. J. Gen. Virol. **70:**3003–3013.
47. **Honess, R. W., U. A. Gompels, B. G. Barrell, M. Craxton, K. R. Cameron, R. Staden, Y.-N. Chang, and G. S. Hayward.** 1989. Deviations from expected frequencies of CpG dinucleotides in herpesvirus DNAs may be diagnostic in the states of their latent genomes. J. Gen. Virol. **70:**837–855.
48. **Jackson, T. R., L. A. C. Blair, J. Marshall, M. Goedert, and M. R. Hanley.** 1988. The *mas* oncogene encodes an angiotensin receptor. Nature (London) **335:**437–440.
49. **Julius, D., T. J. Livelli, T. M. Jessell, and R. Axel.** 1989. Ectopic expression of the serotonin Ic receptor and the triggering of malignant transformation. Science **244:**1057–1062.
50. **Jung, J. U., and R. C. Desrosiers.** 1991. Identification and characterization of the herpesvirus saimiri oncoprotein STP-C488. J. Virol. **65:**6953–6960.
51. **Jung, J. U., J. J. Trimble, N. W. King, B. Biesinger, B. W. Fleckenstein, and R. C. Desrosiers.** 1991. Identification of transforming genes of subgroup A and C strains of herpesvirus saimiri. Proc. Natl. Acad. Sci. USA **88:**7051–7055.
52. **Kamps, M. P., S. S. Taylor, and B. M. Sefton.** 1984. Direct evidence that oncogenic tyrosine kinases and cyclic AMP-dependent protein kinase have homologous ATP-binding sites. Nature (London) **310:**589–592.
53. **Kintner, C., and B. Sugden.** 1981. Conservation and progressive methylation of Epstein-Barr viral DNA sequences in transformed cells. J. Virol. **38:**305–316.
54. **Klickstein, L. B., W. W. Wong, J. A. Smith, J. H. Weis, J. G. Wilson, and D. T. Fearon.** 1987. Human C3b/C4b receptor (CR1). Demonstration of long homologous repeating domains that are composed of the short consensus repeats characteristic of C3/C4 binding proteins. J. Exp. Med. **165:**1095–1112.
55. **Knust, E., S. Schirm, W. Dietrich, W. Bodemer, E. Kolb, and B. Fleckenstein.** 1983. Cloning of herpesvirus saimiri DNA fragments representing the entire L-region of the genome. Gene **25:**281–289.
56. **Koomey, J. M., C. Mulder, R. L. Burghoff, B. Fleckenstein, and R. C. Desrosiers.** 1984. Deletion of DNA sequences in a nononcogenic variant of herpesvirus saimiri. J. Virol. **50:**662–665.
57. **Kotwal, G. J., S. N. Isaacs, R. McKenzie, M. M. Frank, and B. Moss.** 1990. Inhibition of the complement cascade by the major secretory protein of vaccinia virus. Science **250:**827–830.
58. **Kotwal, G. J., and B. Moss.** 1988. Vaccinia virus encodes a secretory polypeptide structurally related to complement control proteins. Nature (London) **335:**176–178.
59. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12:**857–872.
60. **Lachmann, P. J.** 1991. The control of homologous lysis. Immunol. Today **12:**312–315.
61. **Larocca, D., and W. Clough.** 1982. Hypomethylation of Epstein-Barr virus DNA in the nonproducer B-cell line EBR. J. Virol. **43:**1129–1131.
62. **Lauster, R., T. A. Trautner, and M. Noyer-Weidner.** 1989. Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. J. Mol. Biol. **206:**305–312.
63. **Laux, G., M. Perricaudet, and P. J. Farrell.** 1988. A spliced Epstein-Barr virus gene expressed in immortalized lymphocytes is created by circularization of the linear viral genome. EMBO J. **7:**769–774.
64. **Lawrence, G. L., M. Chee, M. A. Craxton, U. A. Gompels, R. W. Honess, and B. G. Barrell.** 1990. Human herpesvirus 6 is closely related to human cytomegalovirus. J. Virol. **64:**287–299.
65. **Lee, S. I., S. C. S. Murthy, J. J. Trimble, R. C. Desrosiers, and J. A. Steitz.** 1988. Four novel U RNAs are encoded by a herpesvirus. Cell **54:**599–607.
66. **Lehner, R., H. Meyer, and M. Mach.** 1989. Identification and characterization of a human cytomegalovirus gene coding for a membrane protein that is conserved among human herpesviruses. J. Virol. **63:**3792–3800.
67. **Lublin, D. M., M. K. Liszewski, T. W. Post, M. A. Arce, M. M. LeBeau, M. B. Rebentisch, R. S. Lemons, T. Seya, and J. P. Atkinson.** 1988. Molecular cloning and chromosomal localization of human membrane cofactor protein (MCP). Evidence for inclusion in the multi-gene family of complement-regulatory proteins. J. Exp. Med. **168:**181–194.
68. **Matsushime, H., M. F. Roussel, R. A. Ashmun, and C. J. Sherr.** 1991. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. Cell **65:**701–713.
69. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69:**1531–1574.
70. **McGeoch, D. J., A. Dolan, S. Donald, and F. J. Rixon.** 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J. Mol. Biol. **181:**1–13.
71. **McGeoch, D. J., A. Dolan, and M. C. Frame.** 1986. DNA sequence of the region in the genome of herpes simplex virus type 1 containing the exonuclease gene and neighboring genes.

Nucleic Acids Res. **14**:3435–3448.
72. **Medof, M. E., D. M. Lublin, V. M. Holers, D. J. Ayers, R. R. Getty, J. F. Leykam, J. P. Atkinson, and M. L. Tykocinski.** 1987. Cloning and characterization of cDNAs encoding the complete sequence of decay-accelerating factor of human complement. Proc. Natl. Acad. Sci. USA **84**:2007–2011.
73. **Medveczky, M. M., P. Geck, C. Clarke, J. Byrnes, J. L. Sullivan, and P. G. Medveczky.** 1989. Arrangement of repetitive sequences in the genome of herpesvirus sylvilagus. J. Virol. **63**:1010–1014.
74. **Medveczky, P., E. Szomolanyi, R. C. Desrosiers, and C. Mulder.** 1984. Classification of herpesvirus saimiri into three groups based on extreme variation in a DNA region required for oncogenicity. J. Virol. **52**:938–944.
75. **Motokura, T., T. Bloom, H. G. Kim, H. Jüppner, J. V. Ruderman, H. M. Kronenberg, and A. Arnold.** 1991. A novel cyclin encoded by a *bcl1*-linked candidate oncogene. Nature (London) **350**:512–515.
76. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
77. **Murphy, P. M., and H. L. Tiffany.** 1991. Cloning of complementary DNA encoding a functional human interleukin-8 receptor. Science **253**:1280–1283.
78. **Murthy, S., J. Kamine, and R. C. Desrosiers.** 1986. Viral-encoded small RNAs in herpes virus saimiri induced tumors. EMBO J. **5**:1625–1632.
79. **Murthy, S. C. S., J. J. Trimble, and R. C. Desrosiers.** 1989. Deletion mutants of herpesvirus saimiri define an open reading frame necessary for transformation. J. Virol. **63**:3307–3314.
80. **Myer, V. E., S. I. Lee, and J. A. Steitz.** 1992. Viral snRNPs bind a protein implicated in messenger RNA destabilization. Proc. Natl. Acad. Sci. USA **89**:1296–1300.
81. **Nicholas, J., K. R. Cameron, H. Coleman, C. Newman, and R. W. Honess.** Analysis of nucleotide sequence of the rightmost 43kbp of herpesvirus saimiri (HVS) L-DNA: general conservation of genetic organization between HVS and Epstein-Barr virus. Virology, in press.
82. **Nicholas, J., K. R. Cameron, and R. W. Honess.** 1992. Herpesvirus saimiri encodes homologous of G protein-coupled receptors and cyclins. Nature (London) **355**:362–365.
83. **Nicholas, J., L. S. Coles, C. Newman, and R. W. Honess.** 1991. Regulation of the herpesvirus saimiri (HVS) delayed-early 110-kilodalton promoter by HVS immediate-early gene products and a homolog of the Epstein-Barr virus R *trans* activator. J. Virol. **65**:2457–2466.
84. **Nicholas, J., U. A. Gompels, M. A. Craxton, and R. W. Honess.** 1988. Conservation of sequence and function between the product of the 52-kilodalton immediate-early gene of herpesvirus saimiri and the BMLF1-encoded transcriptional effector (EB2) of Epstein-Barr virus. J. Virol. **62**:3250–3257.
85. **Nicholas, J., E. P. Smith, L. Coles, and R. Honess.** 1990. Gene expression in cells infected with gammaherpesvirus saimiri: properties of transcripts from two immediate-early genes. Virology **179**:189–200.
86. **Nikas, I., J. McLauchlan, A. J. Davison, W. R. Taylor, and J. B. Clements.** 1986. Structural features of ribonucleotide reductase. Proteins **1**:376–384.
87. **Okada, H., Y. Nagami, K. Takahashi, N. Okada, T. Hideshima, H. Takizawa, and J. Kondo.** 1989. 20 KDa homologous restriction factor of complement resembles T cell activating protein. Biochem. Biophys. Res. Commun. **162**:1553–1559.
88. **Olsen, L. C., R. Aasland, C. U. Wittwer, H. E. Krokan, and D. E. Helland.** 1989. Molecular cloning of human uracil-DNA glycosylase, a highly conserved DNA repair enzyme. EMBO J. **8**:3121–3125.
89. **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.
90. **Pellett, P. E., F. J. Jenkins, M. Ackermann, M. Sarmiento, and B. Roizman.** 1986. Transcription initiation sites and nucleotide sequence of a herpes simplex virus 1 gene conserved in the Epstein-Barr virus genome and reported to affect the transport of viral glycoproteins. J. Virol. **60**:1134–1140.
91. **Preston, V. G., and F. B. Fisher.** 1984. Identification of the herpes simplex virus type 1 gene encoding the dUTPase. Virology **138**:58–68.
92. **Reid, K. B. M., and A. J. Day.** 1989. Structure-function relationships of the complement components. Immunol. Today **10**:177–180.
93. **Ripoche, J., A. J. Day, T. J. R. Harris, and R. B. Sim.** 1988. The complete amino acid sequence of human complement factor H. Biochem. J. **249**:593–602.
94. **Roizman, B.** 1982. The family Herpesviridae: general description, taxonomy, and classification, p. 1–23. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Press, New York.
95. **Roizman, B.** 1990. Herpesviridae: a brief introduction, p. 1787–1793. *In* B. N. Fields, D. M. Knipe, et al. (ed.), Virology, 2nd ed. Raven Press Ltd., New York.
96. **Rosa, M. D., E. Gottlieb, M. R. Lerner, and J. A. Steitz.** 1981. Striking similarities are exhibited by two small Epstein-Barr virus-encoded ribonucleic acids and the adenovirus-associated ribonucleic acids VAI and VAII. Mol. Cell. Biol. **1**:785–796.
97. **Rosenberg, C. L., E. Wong, E. M. Petty, A. E. Bale, Y. Tsujimoto, N. L. Harris, and A. Arnold.** 1991. PRAD1, a candidate *BCL1* oncogene: mapping and expression in centrocytic lymphoma. Proc. Natl. Acad. Sci. USA **88**:9638–9642.
98. **Sample, J., D. Liebowitz, and E. Kieff.** 1989. Two related Epstein-Barr virus membrane proteins are encoded by separate genes. J. Virol. **63**:933–937.
99. **Sawada, R., K. Ohashi, H. Anaguchi, H. Okazaki, M. Hattori, N. Minato, and M. Naruto.** 1990. Isolation and expression of the full-length cDNA encoding CD59 antigen of human lymphocytes. DNA Cell Biol. **9**:213–220.
100. **Schirm, S., I. Müller, R. C. Desrosiers, and B. Fleckenstein.** 1984. Herpesvirus saimiri DNA in a lymphoid cell line established by in vitro transformation. J. Virol. **49**:938–946.
101. **Stamminger, T., R. W. Honess, D. F. Young, W. Bodemer, E. D. Blair, and B. Fleckenstein.** 1987. Organization of terminal reiterations in the virion DNA of herpesvirus saimiri. J. Gen. Virol. **68**:1049–1066.
102. **Storz, J., B. Ehlers, W. J. Todd, and H. Ludwig.** 1984. Bovine cytomegaloviruses: identification and differential properties. J. Gen. Virol. **65**:697–706.
103. **Stow, N. D., and E. C. McMonagle.** 1983. Characterization of the TR_s/IR_s origin of DNA replication of herpes simplex virus type 1. Virology **130**:427–438.
104. **Sugita, Y., T. Tobe, E. Oda, M. Tomita, K. Yasukawa, N. Yamaji, T. Takemoto, K. Furuichi, M. Takayama, and S. Yano.** 1989. Molecular cloning and characterization of MACIF, an inhibitor of membrane channel formation of complement. J. Biochem. **106**:555–557.
105. **Szomolanyi, E., P. Medveczky, and C. Mulder.** 1987. In vitro immortalization of marmoset cells with three subgroups of herpesvirus saimiri. J. Virol. **61**:3485–3490.
106. **Thompson, R., R. W. Honess, L. Taylor, J. Morran, and A. J. Davison.** 1987. Varicella-zoster virus species a thymidylate synthetase. J. Gen. Virol. **68**:1449–1455.
107. **Thomson, B. J., and J. Nicholas.** 1991. Superantigen function. Nature (London) **351**:530.
108. **Trimble, J. J., S. C. S. Murthy, A. Bakker, R. Grassmann, and R. C. Desrosiers.** 1988. A gene for dihydrofolate reductase in a herpesvirus. Science **239**:1145–1147.
109. **Walker, J. E., M. Saraste, M. J. Runswick, and N. J. Gay.** 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO J. **1**:945–951.
110. **Wassarman, D. A., S. I. Lee, and J. A. Steitz.** 1989. The nucleotide sequence of HSUR 5 RNA from herpesvirus saimiri. Nucleic Acids Res. **17**:1258.
111. **Weis, J. J., L. E. Toothaker, J. A. Smith, J. H. Weis, and D. T. Fearon.** 1988. Structure of the human B lymphocyte receptor for C3d and the Epstein-Barr virus and relatedness to other members of the family of C3/C4 binding proteins. J. Exp. Med. **167**:1047–1066.

112. **Welch, A. R., L. M. McNally, and W. Gibson.** 1991. Cytomegalovirus assembly protein nested gene family: four 3′-coterminal transcripts encode four in-frame, overlapping proteins. J. Virol. **65**:4091–4100.
113. **Welch, A. R., A. S. Woods, L. M. McNally, R. J. Cotter, and W. Gibson.** 1991. A herpesvirus maturational proteinase, assemblin: identification of its gene, putative active site domain, and cleavage site. Proc. Natl. Acad. Sci. USA **88**:10792–10796.
114. **Wittenberg, C., K. Sugimoto, and S. I. Reed.** 1990. G1-specific cyclins of S. cerevisiae: cell cycle periodicity, regulation by mating pheromone, and association with the $p34^{CDC28}$ protein kinase. Cell **62**:225–237.
115. **Xiong, Y., T. Connolly, B. Futcher, and D. Beach.** 1991. Human D-type cyclin. Cell **65**:691–699.
116. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A *cis*-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc. Natl. Acad. Sci. USA **81**:3806–3810.
117. **Zhu, L., and S. K. Weller.** 1992. The six conserved helicase motifs of the UL5 gene product, a component of the herpes simplex virus type 1 helicase-primase, are essential for its function. J. Virol. **66**:469–479.