JOURNAL OF VIROLOGY, Jan. 1990, p. 287–299
0022-538X/90/010287-13$02.00/0

Vol. 64, No. 1

Human Herpesvirus 6 Is Closely Related to Human Cytomegalovirus

G. L. LAWRENCE,[1]* M. CHEE,[1] M. A. CRAXTON,[2] U. A. GOMPELS,[2] R. W. HONESS,[2] AND B. G. BARRELL[1]

Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH,[1] and Division of Virology, National Institute for Medical Research, London NW7 1AA,[2] United Kingdom

Received 29 June 1989/Accepted 25 September 1989

A sequence of 21,858 base pairs from the genome of human herpesvirus 6 (HHV-6) strain U1102 is presented. The sequence has a mean composition of 41% G+C, and the observed frequency of CpG dinucleotides is close to that predicted from this mononucleotide composition. The sequence contains 17 complete open reading frames (ORFs) and part of another at the 5′ end of the sequence. The predicted protein products of two of these ORFs have no recognizable homologs in the genomes of other sequenced human herpesviruses (i.e., Epstein-Barr virus [EBV], human cytomegalovirus [HCMV], herpes simplex virus [HSV], and varicella-zoster virus [VZV]). However, the products of nine other ORFs are clearly homologous to a set of genes that is conserved in all other sequenced herpesviruses, including homologs of the alkaline exonuclease, the phosphotransferase, the spliced ORF, and the major capsid protein genes. Measurements of similarity between these homologous sequences showed that HHV-6 is clearly most closely related to HCMV. The degree of relatedness between HHV-6 and HCMV was commensurate with that observed in comparisons between HSV and VZV or EBV and herpesvirus saimiri and significantly greater than its relatedness to EBV, HSV, or VZV. In addition, the gene for the major capsid protein and its 5′ neighbor are reoriented with respect to the spliced ORFs in the genomes of both HHV-6 and HCMV relative to the organization observed in EBV, HSV, and VZV. Three ORFs in HHV-6 have recognizable homologs only in the genome of HCMV. Despite differences in gross composition and size, we conclude that the genomes of HHV-6 and HCMV are closely related.

The first recognized isolations of a previously undetected human herpesvirus, now called human herpesvirus 6 (HHV-6), were obtained in the course of in vitro cultivation of peripheral blood lymphocytes from patients with lymphoproliferative disorders, some of whom were also infected with human immunodeficiency virus (50). The viruses were shown to have the ultrastructural and morphogenetic properties characteristic of a herpesvirus (5, 59) but to be distinct from the five previously known human herpesviruses by their antigenic properties and by the failure to show homologous hybridization with nucleic acid sequences from each of these other five human viruses (32). Independent isolates of herpesviruses shown to be closely related to the initial isolate (HBLV/GS) were subsequently reported from human immunodeficiency virus-infected patients from Uganda (strains U1102 and U683 [20]), The Gambia (strain AJ [59]), and Zaire (strain Z29 [38]). A series of seroepidemiological investigations has since established that evidence of a prior infection with HHV-6 is widespread in populations of apparently healthy adults and that the virus is typically acquired in early infancy (7, 52). The primary infection in infants has been shown to cause the common childhood infection exanthem subitum (roseola infantum [34, 61]), and a series of virus isolations from the acute stages of this mild childhood disease has been obtained. There have also been reports of the common detection of HHV-6 in cervical lymph nodes (23) and of HHV-6 DNA sequences in a proportion of some rare B-cell tumors (31) and suggestions that infection or recurrence in adult life may be related to lymphadenopathy (9, 46).

We are interested in the relationships between the divergent biological and molecular genetic properties of the herpesviruses and their evolution. The current classification of the herpesviruses recognizes their biological diversity and divides them into three subgroups (the alpha-, beta-, and gammaherpesviruses) on the basis of some of these biological properties (28, 30, 49). Alphaherpesviruses, exemplified by herpes simplex viruses (herpes simplex virus types 1 and 2 [HSV-1 and HSV-2]; human herpesviruses 1 and 2) and varicella-zoster virus (VZV; human herpesvirus 3), are distinguished by their capacity to establish latent infections of neural tissues and to reactivate from these sites. Betaherpesviruses include the cytomegaloviruses (e.g., human cytomegalovirus [HCMV]; human herpesvirus 5); they replicate productively in cultures of fibroblasts from the host species. The sites of their persistence in vivo are uncertain but may involve reticuloendothelial cells and do not appear to involve neural tissues (42). Gammaherpesviruses are typified by the B-cell lymphotropic human herpesvirus, Epstein-Barr virus (EBV; human herpesvirus 4), and the T-cell lymphotropic virus of the squirrel monkey, herpesvirus saimiri (HVS; saimiriine herpesvirus 2). The major mode of virus persistence of these lymphotropic viruses is as latent infections of circulating lymphocytes. The isolations of HHV-6 from peripheral blood lymphocytes have clearly shown that the virus can infect a population of lymphocytes in vivo. The major population of productively infected cells in cultures of cord blood or peripheral blood lymphocytes has the characteristics of immature $CD4^+$ T cells (20, 39), and the virus can be propagated in cultures of lymphoblastoid cells in vitro (39, 59). Despite the lack of knowledge on the nature of the latent site of the virus or any demonstration that HHV-6 can transform lymphoid cells, it has been suggested that the virus should provisionally be classified as a gammaherpesvirus (33, 38).

We have undertaken an analysis of the structure and sequence of the genome from a Ugandan isolate of HHV-6 (U1102 [20]). In this report, we present and interpret a

* Corresponding author.

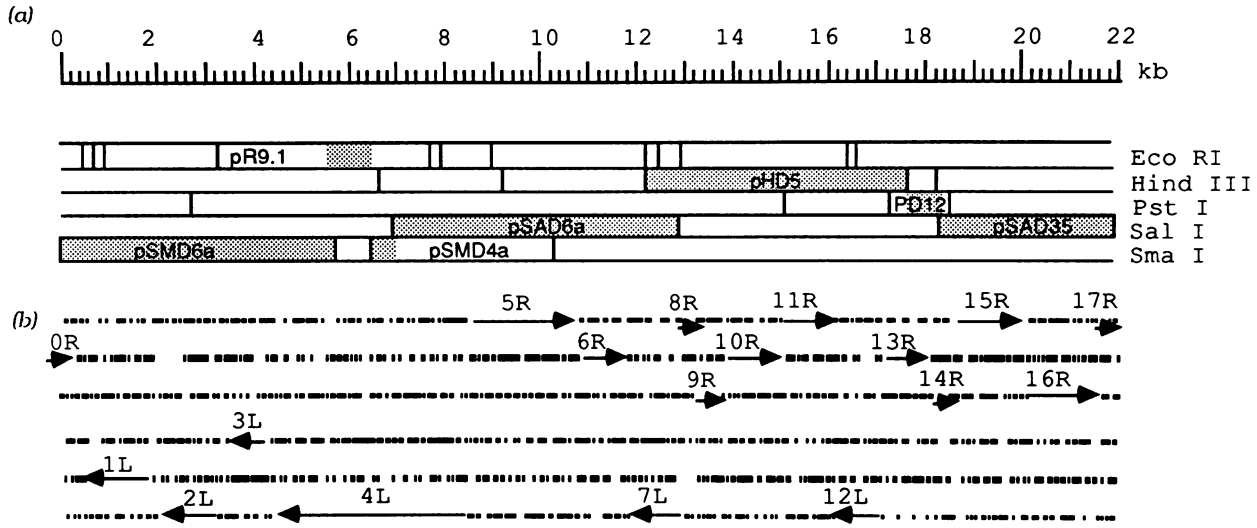


FIG. 1. (a) Restriction map of the region sequenced showing *Eco*RI, *Hin*dIII, *Pst*I, *Sma*I, and *Sal*I restriction sites. The DNA sequence was determined from the shaded portions of the plasmids, whose names appear on the restriction fragments from which they were derived. (b) Positions of termination codons in each of the three possible ORFs for each strand of the region sequenced. Arrows indicate the location and direction of the major ORFs. ORFs are named 0 to 17; R or L indicates rightward or leftward orientation.

sequence of 21.8 kilobase pairs (kbp) from the genome of HHV-6 that includes the sequence previously recognized as having significant nucleotide sequence similarities to a region of the HCMV genome (22). The sequence and arrangement of the predicted open reading frames (ORFs) in this region of the HHV-6 genome bear a much closer resemblance to corresponding regions of HCMV (a betaherpesvirus) than to corresponding regions of the genome of EBV, HSV, or VZV.

MATERIALS AND METHODS

Isolation and characterization of recombinant DNA clones of HHV-6 DNA. All recombinant DNA clones were isolated from HHV-6 (U1102) DNA prepared from cultures of infected cord blood lymphocytes. The *Sal*I and *Sma*I clones were prepared by cloning purified restriction endonuclease fragments into the *Sal*I site of a pBS (Bluescribe; Stratagene) vector or the *Sma*I site of pUC13. Fragments were selected for sequencing on the basis of their linkage relationships to the 5.4-kbp *Hin*dIII fragment (cloned into the *Hin*dIII site of pUC8 as pHD5), which was previously shown to be homologous to a region of the HCMV genome (22; unpublished results). The *Eco*RI plasmid pR9.1 was provided by M. Jones, and the PD12 clone was a 1.2-kbp *Pst*I fragment cloned directly into the *Pst*I site of M13mp18. The relationships between these cloned fragments over the relevant portion of the HHV-6 genome are summarized in Fig. 1. A detailed description of the mapping and cloning of the HHV-6 (U1102) genome will be presented elsewhere, but the rightmost *Sal*I site of pSAD3.5 in the region analyzed in this report is located approximately 31 kbp from the right unique/repeat junction of the HHV-6 genome, the total size of which is about 170 kbp.

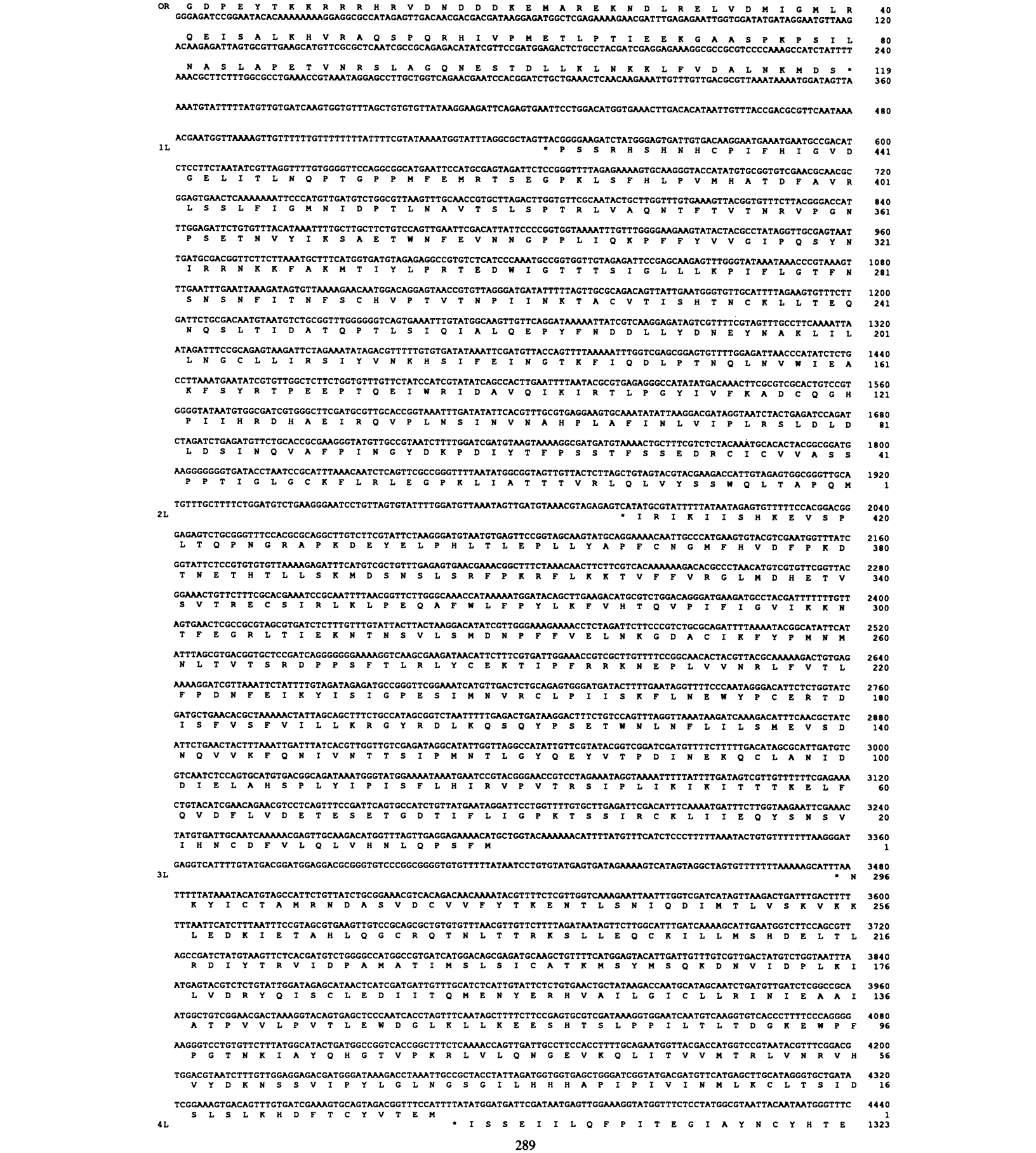
DNA sequencing and sequence analysis. The DNA sequence was determined from the regions shown in Fig. 1, using the methods described by Bankier et al. (2). Random subfragments of DNA from these plasmids were prepared by sonication (19) and subcloned into M13mp8 (43), and single-stranded templates were sequenced by the dideoxynucleotide-chain termination method (51). Regions of sequence compression were resolved by replacing dGTP with deoxy-7-deazaguanosine triphosphate in the sequencing reactions (45).

Sequence data were assembled by using the computer programs DBAUTO and DBUTIL (55, 56) and analyzed for the presence of ORFs and transcription signals with the programs DIANA (J. Crooke, T. S. Horsnell, and B. G. Barrell, unpublished data) and ANALYSEQ (58). Predicted protein sequences were analyzed for hydrophobicity and potential glycosylation sites with ANALYSEP (58), and searches for homologous protein sequences contained in protein libraries were performed by using the computer program FASTP (37). The AMPS suite of programs (3, 4) was used to carry out pairwise computer alignments of predicted translation products of HHV-6 ORFs and the homologous genes of the other human herpesviruses. Twenty randomizations of each alignment were performed so that a significance score for the alignment could be obtained. The program uses the Dayhoff mutation data matrix (17, 18) for protein alignments. All computer programs were run on DEC VAX and microVAX computers.

RESULTS

The DNA sequence of a 21,858-bp region of the HHV-6 (U1102) genome has been determined for both strands, each base being sequenced an average of six times by the random

FIG. 2. DNA and predicted protein sequences. The nucleotide sequence reported here will appear in the EMBL and GenBank data bases under accession number M28243. The DNA sequence is given as the rightward 5′-to-3′ strand only (numbered 1 to 21858). Rightward-encoded protein sequences are shown above the corresponding DNA sequences in single-letter code; leftward-encoded protein sequences are shown below the corresponding DNA sequences. The name of each ORF is given on the left of the first line of sequence, and amino acid sequences are numbered from the N terminus to the C terminus to the right of the sequence. Protein sequences are shown from the first ATG. The sequence continues on the following pages.



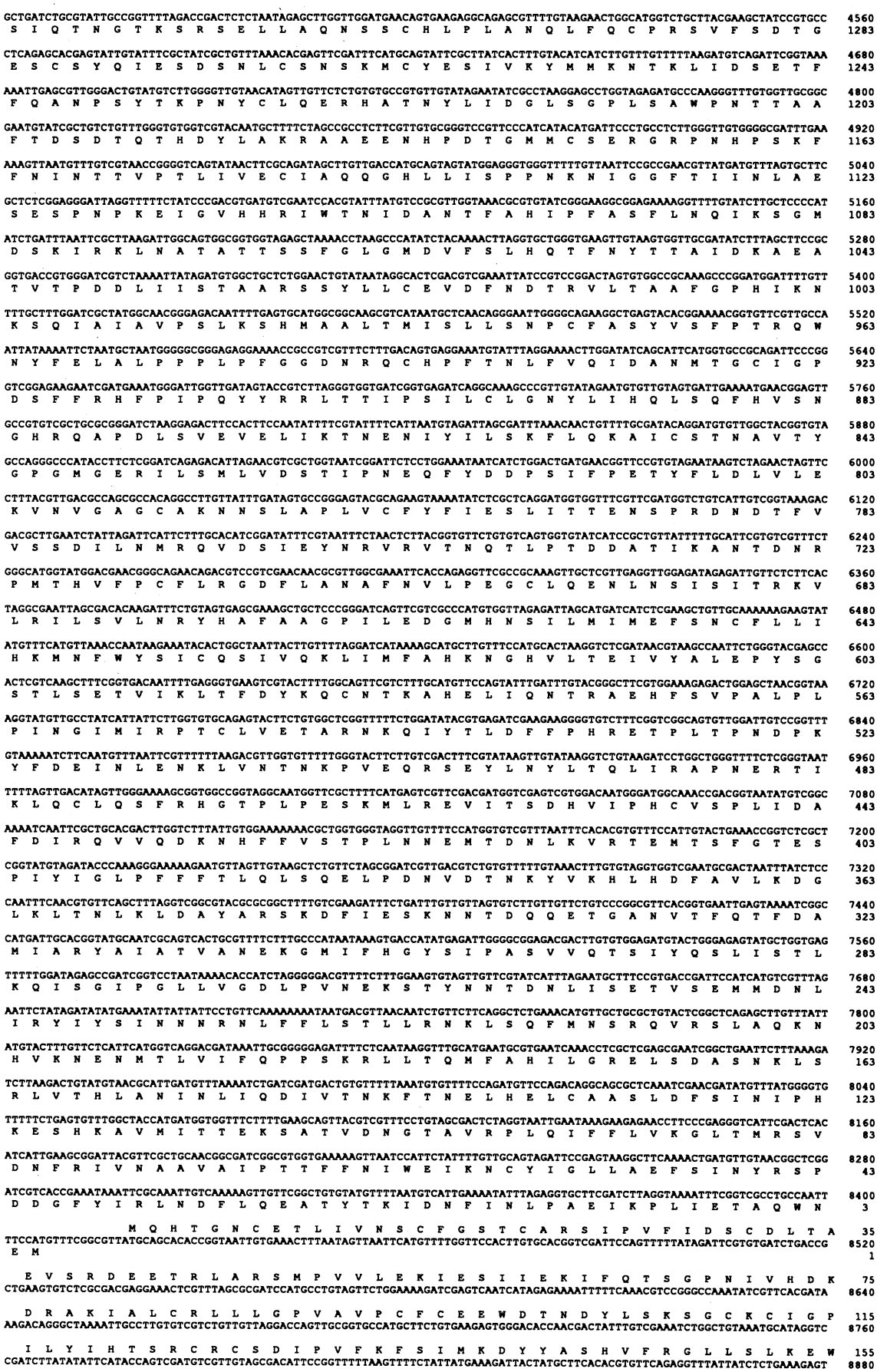


FIG. 2—*Continued.*

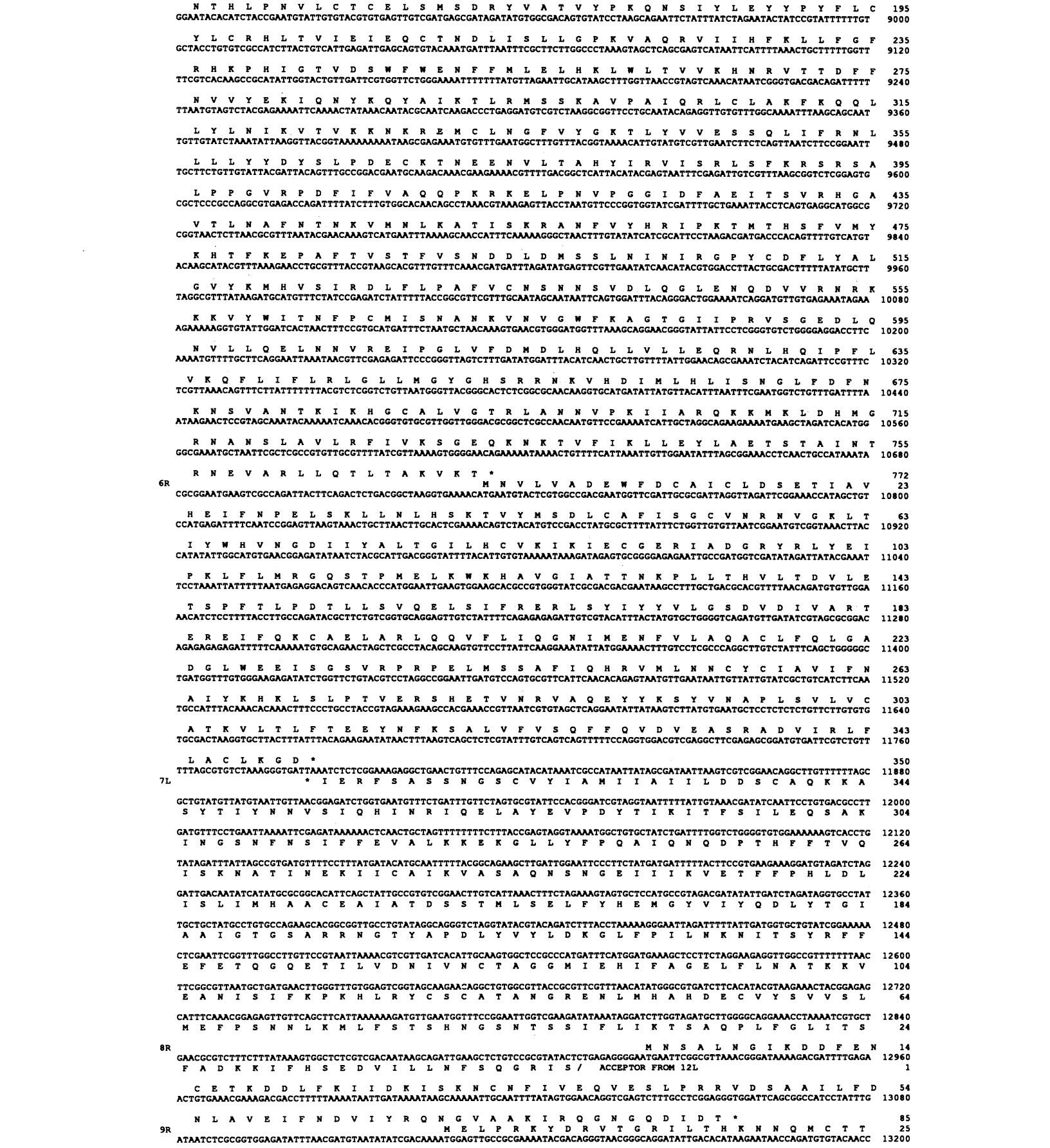


FIG. 2—*Continued.*

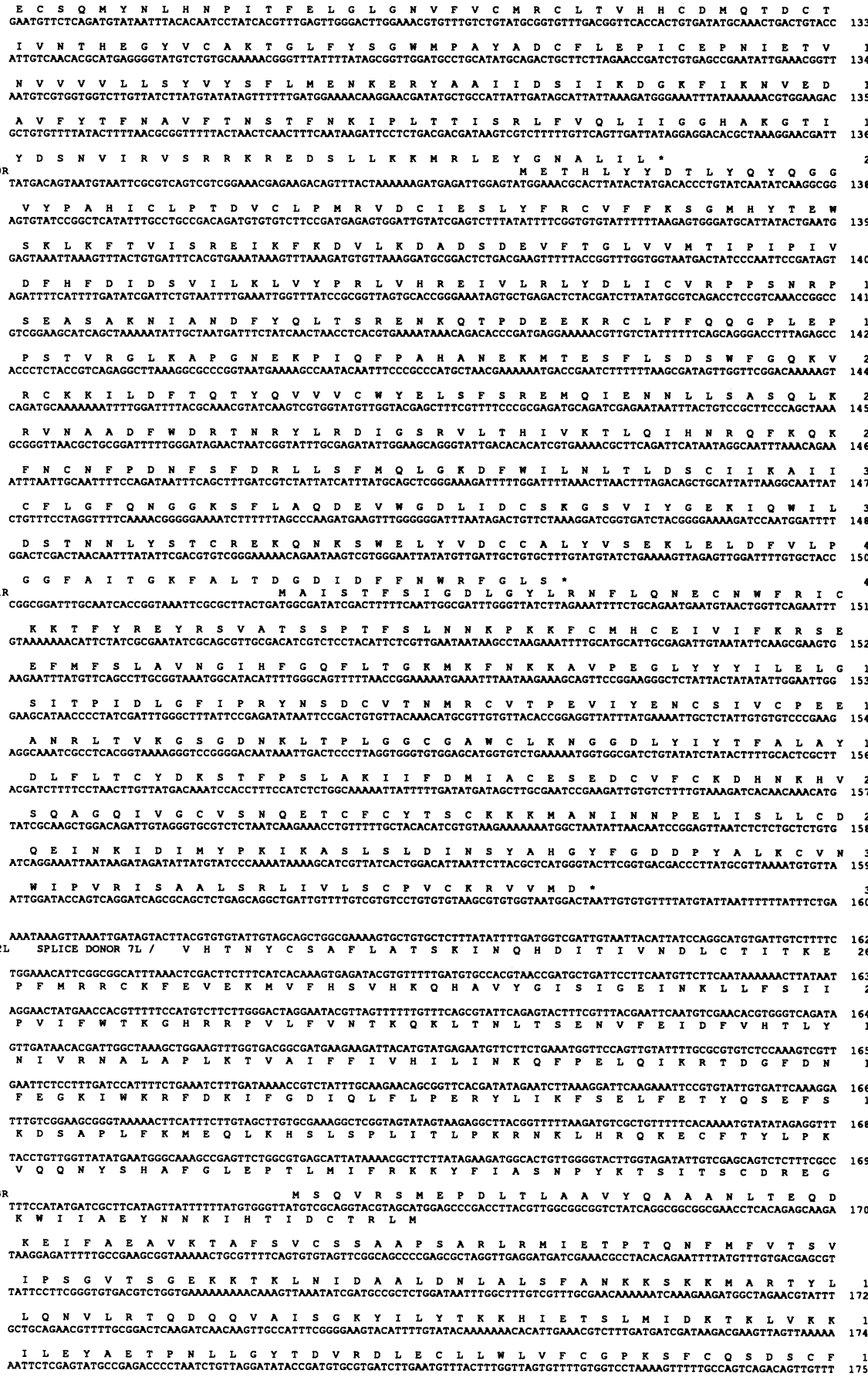


FIG. 2—*Continued.*

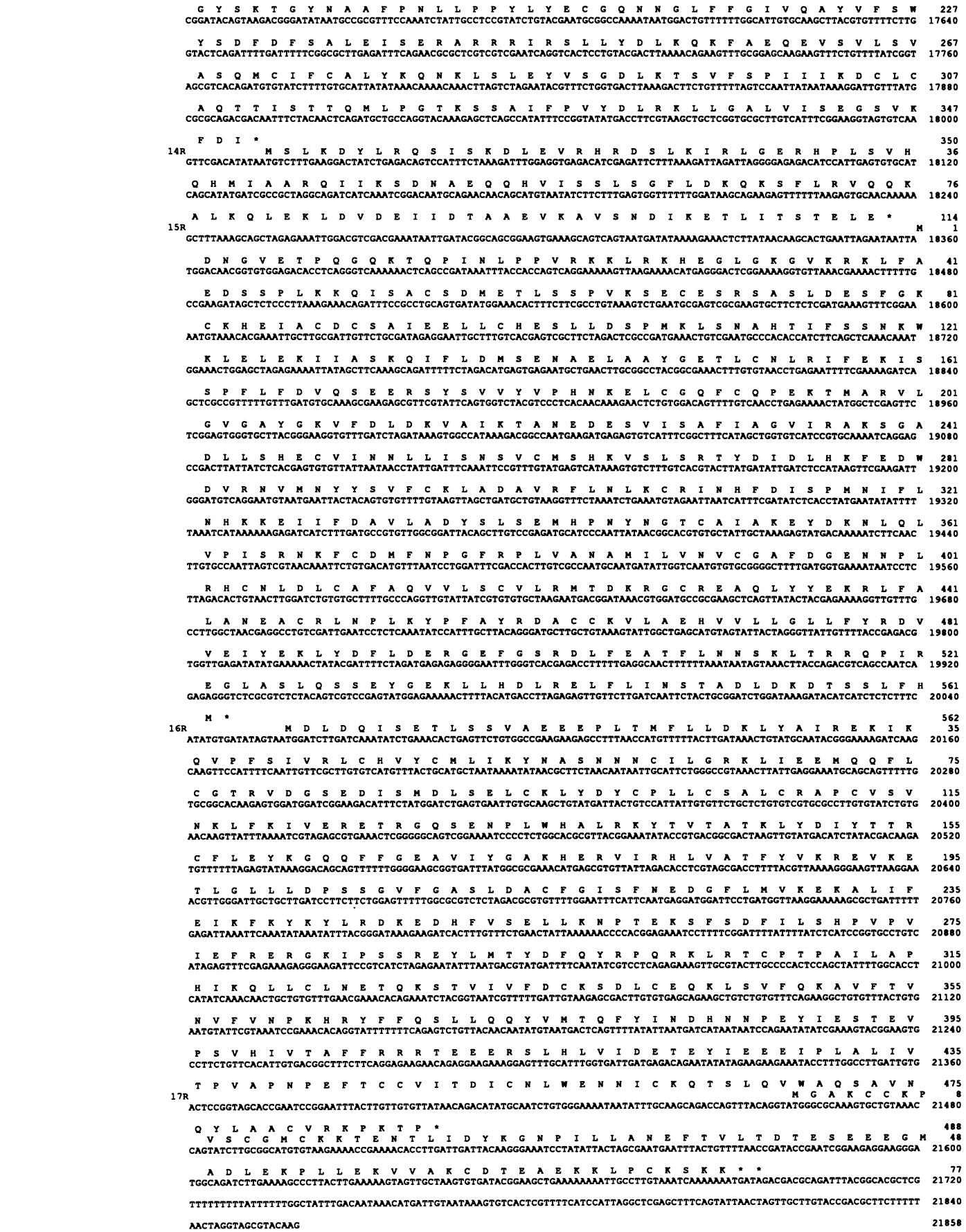


FIG. 2—*Continued*.

TABLE 1. Summary of data: ORFs, putative translation start sites, TATA consensus sequences, and lengths and relative molecular masses of predicted translation products

Name	ORF start	ORF end	ATG position	ATG context sequence	TATA position	TATA sequence	Length (amino acids)	Molecular size (kilodaltons)
0R	(2)[a]	358					(119)[a]	
1L	1954	544	1921	AACATGC	1978	TATTTAA	458	51.5
2L	3315	2002	3298	AGCATGT	3330	CATAAAA	432	50.2
					3341	TATTTAA		
3L	4412	3480	4367	AAAATGG	?[b]	?	296	33.5
4L	8415	4372	8406	AACATGG	8450	TATTAAA	1,345	151.9
					8501	TATAAAA[c]		
5R	8376	10733	8418	GTTATGC	8352	TATTTAGA[c]	772	88.7
6R	10724	11782	10733	AACATGG	10673	CATAAATA	350	39.9
7L	12984	11785						
	12909[d]						666	76.2
12L	17007	16066	16989	TCCATGC[e]	17146	CATAAAA		
		16108[f]						
8R	12891	13175	12921	GGAATGA	12856	TATAAAG	85	9.6
9R	13114	13773	13126	AAAATGG	13099	TATTTAA	216	24.9
10R	13745	15079	13754	AGTATGG	13524	TATTAAA[c]	432	51.4
					13542	TATAAAA[c]		
11R	14948	16043	15039	CTGATGG	?	?	335	37.9
13R	16943	18009	16952	TTTATGT[e]	16849	TATAAAA[c]	353	39.4
			16961	GTTATGT	16849	TATAAAA[c]		
			16979	AGCATGG	16944	TATTTTT		
14R	17962	18354	18013	ATAATGT	17861	TATAATA[c]	114	13.2
15R	18348	20045	18360	ATTATGG	18318	TATAATA[c]	561	63.7
16R	19996	21519	20056	GTAATGG	20016	GATAAAG[c]	488	56.6
17R	21432	21689	21459	GGTATGG	21398	TATAACA	77	8.3

[a] Incomplete ORF.
[b] ?, No obvious TATA consensus 5′ to the first ATG of the ORF.
[c] An intervening ATG codon lies between the proposed TATA consensus sequence and the first ATG of the ORF.
[d] Splice acceptor site of ORF 7L.
[e] ATG context sequence does not conform to the Kozak (36) consensus sequence (RNNATG or NNNATGG).
[f] Splice donor site of ORF 12L.

method. The overall G+C content was found to be 41%, considerably lower than the values for HSV-1 (68%) (40), EBV (60%) (1), and HCMV (58%) (60; M. S. Chee, 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, Curr. Top. Microbiol. Immunol., in press) and in the same range as those for VZV (46%) (15) and HVS L DNA (36%) (25). Observed frequencies of CpG dinucleotides in this portion of the HHV-6 DNA sequence did not differ significantly from those expected on the basis of random associations between mononucleotides (not shown). In contrast, eucaryotic DNA (6) and the genomes of the gammaherpesviruses EBV and HVS do show an overall CpG dinucleotide deficiency (27, 29).

The region sequenced contains 1 partial and 17 complete ORFs, numbered 0R (R, rightward; L, leftward) to 17R in Fig. 1, which were predicted to be coding by the positional base preferences method of Staden (57) (data not shown). The protein sequences of the predicted ORFs (beginning from the first ATG codon) are shown relative to the nucleotide sequence in Fig. 2. A summary of the proposed locations of the ORFs together with positions of putative TATA boxes and translation start sites appears in Table 1. Potential polyadenylation signals (AATAAA and ATTAAA) encoded by the viral genome are numerous, presumably because of the high A+T content of the DNA sequence. No attempt has been made to predict which ones may be used during transcription of viral genes because no transcription mapping data are available.

TABLE 2. Summary of optimized FASTP scores observed in comparisons between HHV-6 ORF 11R and the homologous genes from the other human herpesviruses

Virus	ORF	Score				
		11R	HCMVUL94	BGLF2	UL16	44
HHV-6	11R	1821				
HCMV	HCMVUL94	530	1929			
EBV	BGLF2	221	236	1718		
HSV-1	UL16	108	62	169	1986	
VZV	44	<30	<30	<30	518	1857

Homology of HHV-6 ORFs to those of other herpesviruses. Each of the 18 amino acid sequences was screened against a library of herpesvirus protein sequences as well as the Protein Information Resource library (26), using the computer program FASTP (37) with a K-tuple value of 2. An optimized FASTP score of greater than 100 was considered to indicate a significant degree of amino acid similarity. No nonherpesvirus proteins were identified when the Protein Information Resource library was searched. If a homologous sequence in the herpesvirus protein library was identified, then it was used to rescreen the library. Thus, the evolutionary relationships of proteins from other herpesviruses to those of HHV-6 could be established even if the similarities observed in some direct pairwise comparisons were of uncertain significance. For example, when the HHV-6 ORF 11R amino acid sequence was used as the probe sequence, it identified homologous protein sequences encoded in the

TABLE 3. Summary of identification of homologs of the HHV-6 ORFs in the genomes of other human herpesviruses

HHV-6	HCMV	EBV	HSV-1	VZV	Comment
0R	NS[a]	NS	NS	NS	Incomplete
1L	NS	NS	NS	NS	Potentially glycosylated
2L	NS	NS	NS	NS	Hydrophobic
3L	HCMVUL85	BDLF1	UL18	41	—[b]
4L	HCMVUL86	BcLF1	UL19	40	Major capsid protein
5R	HCMVUL87	BcRF1	NS	NS	Hydrophobic
6R	HCMVUL88	NS	NS	NS	
7L	HCMVUL89EX2	BDRF1	UL15	42	Splice exon 2
NS	HCMVUL90	NS	NS	NS	
8R	HCMVUL91	NS	NS	NS	
9R	HCMVUL92	BDLF4	NS	NS	
10R	HCMVUL93	BGLF1	UL17	43	—[b]
11R	HCMVUL94	BGLF2	UL16	44	—[b]
12L	HCMVUL89EX1	BGRF1	UL15	45	Splice exon 1
13R	HCMVUL95	BGLF3	NS	NS	
14R	HCMVUL96	NS	NS	NS	
15R	HCMVUL97	BGLF4	UL13	47	Putative phoshophotransferase
16R	HCMVUL98	BGLF5	UL12	48	Alkaline exonuclease
17R	NS	BBLF1	UL11	49	Hydrophilic[b]

[a] NS, No significant sequence similarity.
[b] Identified by the cross-referencing method shown in Table 2, whereby each ORF amino acid sequence identified homologs in at least one but not all of the other human herpesviruses during library searches with the program FASTP.

genomes of HCMV, EBV, and HSV-1, but a VZV homolog was identified only when the HSV-1 UL16 amino acid sequence was used as the probe sequence (Table 2). Table 3 lists all of the HHV-6 ORFs found in this study alongside their counterparts in other human herpesviruses.

The region contains ORFs homologous to a number of highly conserved genes, including the major capsid gene (10, 14), the two exons of the spliced gene of unknown function (12), a putative phosphotransferase gene (11, 54), and the alkaline exonuclease gene (21, 41). The organization of these ORFs with respect to size and orientation is most similar to that seen in HCMV (Fig. 3).

Calculations of the percentage amino acid identity shared by HHV-6 ORFs and their counterparts in HCMV, EBV, HSV-1, and VZV by using the AMPS program (3, 4) revealed that the majority of HHV-6 ORFs shared the greatest degree of sequence similarity with their HCMV homologs and least with those of the alphaherpesviruses. A comparison of the percentage amino acid identity shared by HCMV and HHV-6 ORFs and HHV-6 and EBV ORFs is presented in Table 4. Significance scores were calculated from 20 randomizations of each pairwise comparison.

Analysis of ORFs. The first three ORFs, 0R, 1L, and 2L, have no significant amino acid sequence similarity to ORFs in other human herpesviruses. ORF 0R is the C-terminal portion of an ORF that begins in the neighboring restriction fragment. Good candidates for TATA and polyadenylation signals are located close to 5′ and 3′ ends of both 1L and the neighboring ORF 2L (Table 1). ORF 1L encodes a protein product 458 amino acids in length that has eight potential N-linked glycosylation sites but no obvious hydrophobic signal or transmembrane anchor sequences, which would be more clearly indicative of a function as a surface glycoprotein. Other ORFs encode potential N-linked glycosylation

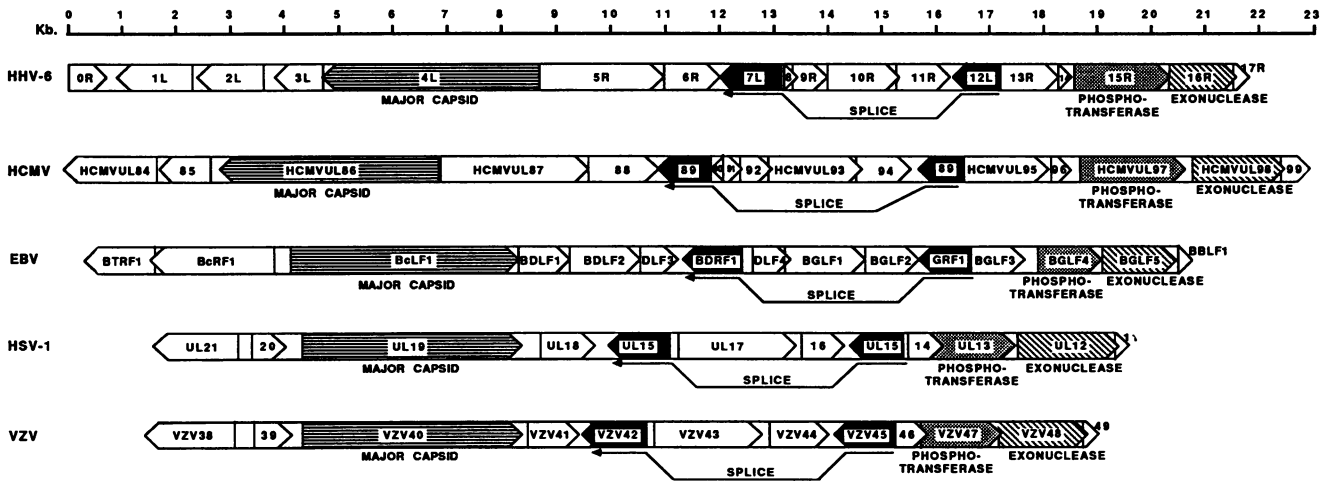


FIG. 3. Arrangement of ORFs of HHV-6 compared with the arrangement of ORFs for the homologous regions of HCMV (Chee et al., in press), EBV (1), HSV-1 (40), and VZV (15). The diagram is oriented so that the exons of the conserved spliced gene lie in the same direction for each of the herpesviruses. The shaded ORFs are those that encode proteins whose amino acid sequence is conserved across each of the herpesviruses represented.

TABLE 4. Percentage of amino acid identity obtained in comparisons between the predicted protein sequences of HHV-6 ORFs and those of the homologous genes of HCMV and EBV

HCMV ORF	% Identity	Significance[a]	HHV-6 ORF	% Identity	Significance	EBV ORF
HCMVUL85	44	56	3L	23	15	BDLF1
HCMVUL86	43	56	4L	29	36	BcLF1
HCMVUL87	43	52	5R	32	16	BcRF1
HCMVUL88	29	17	6R			
HCMVUL89EX2	61	45	7L[b]	42	40	BDRF1
HCMVUL91	31	8	8R			
HCMVUL92	45	17	9R	23	14	BDLF4
HCMVUL93	23	12	10R	21	4	BGLF1
HCMVUL94	33	26	11R	22	14	BGLF2
HCMVUL89EX1	52	28	12L[b]	28	16	BGRF1
HCMVUL95	37	26	13R	18	7	BGLF3
HCMVUL96	25	31	14R			
HCMVUL97	27	18	15R	20	2	BGLF4
HCMVUL98	37	29	16R	27	16	BGLF5
			17R	28	2	BBLF1
HCMVUL89[c]	57	88	12L7L[c]	36	36	BGDRF1[c]

[a] Score indicating the significance of the pairwise alignment on the basis of 20 randomizations of the two sequences. A score of <5 indicates the need for close examination of the alignment, and a score of >15 indicates that the alignment is likely to be meaningful (3).
[b] Exons 1 and 2 of spliced ORF analyzed separately.
[c] Exons 1 and 2 of spliced ORF analyzed together.

sites, but only ORFs 1L, 4L, and 7L have more than three sites (8, 13, and 6, respectively).

The next three ORFs, 3L, 4L, and 5R, are homologous, with a contiguous block of ORFs in both HCMV and EBV. Two of these (3L and 4L) are conserved across all of the known human herpesviruses. This block of genes is inverted in both HHV-6 and HCMV (HCMVUL85 and HCMVUL86) compared with the arrangement observed for the homologous genes in EBV (BDRF1 and BcLF1), HSV-1 (UL18 and UL19), and VZV (VZV41 and VZV40). The HHV-6 major capsid protein is encoded by 4L. It is 1,345 amino acids long, shorter than its homologs in the other human herpesviruses by 25 to 51 amino acids.

ORFs 6R, 8R, and 14R have homologs in HCMV only. The arrangement of these ORFs, as well as 9R, 10R, and 11R, which are located on the opposite strand within the intron of the spliced gene, most closely resembles that observed in HCMV (Fig. 3).

The next two rightward ORFs (15R and 16R) encode highly conserved genes which in EBV are transcribed during the early stages of productive infection (24). ORF 15R encodes a putative phosphotransferase and is discussed in more detail elsewhere (10). The significance score for the pairwise amino acid alignment of ORF 15R and its EBV homolog is <5 (Table 4) and is probably due to conservation of amino acid motifs within a divergent overall amino acid sequence. ORF 16R encodes the HHV-6 homolog of the alkaline exonuclease gene (21, 41). Although this ORF has the highest percentage of identical amino acids with the homologous gene of HCMV, it has a higher overall amino acid similarity to its counterpart in EBV, reflected as a higher optimized FASTP score (409 versus 374).

ORF 17R overlaps 16R by an unusually large proportion of its total length (20 of 76 amino acids). The presence of a proposed TATA box and ATG as well as the positional base preference analysis (not shown) all indicate that the ORF is likely to be a protein-coding sequence. The most similar gene in the other herpesviruses is ORF 49 of VZV, with which it shares 32% identity over a region of 44 amino acids. VZV ORF 49 encodes a hydrophilic protein (15), as does ORF 17R. Although the other human herpesviruses have similar-size ORFs located adjacent to the alkaline exonucle-

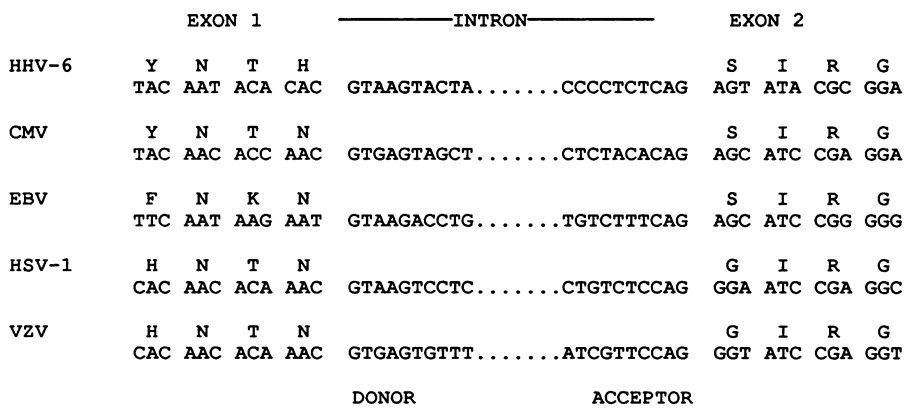


FIG. 4. DNA sequence and encoded protein sequences in the region of the predicted splice donor site of exon 1 and the predicted splice acceptor site of exon 2 of the conserved spliced gene for HHV-6, HCMV (Chee and Barrell, in preparation), EBV (1), HSV-1 (12, 40), and VZV (15). The conserved nucleotides that are found in the splice donor consensus sequence and the splice acceptor consensus sequence are (C/A)AG:GT(A/G)AGT and (C/T)AG:G, respectively, where : represents the intron-exon junction (53).

TABLE 5. Percentage amino acid similarity observed in comparisons between a portion of exon 2 of the conserved spliced gene from representative herpesviruses

Group	Virus	% Amino acid similarity					
		HHV-6	HCMV	EBV	HVS	HSV-1	VZV
α	VZV	46	48	41	27	68	100
α	HSV-1	47	47	46	34	100	
γ	HVS	44	36	68	100		
γ	EBV	44	49	100			
β	HCMV	66	100				
	HHV-6	100					
	% G+C of corresponding genome	43	57	60	36	68	46

ase gene (the UL11 ORF of HSV-1 [40] and the BBLF1 ORF of EBV, which is transcribed at late times in productive viral infection [24]), their similarities to HHV-6 17R are not sufficient to give evidence of homology. The ORF in the comparable position in the HCMV sequence encodes the highly immunogenic 28-kilodalton phosphoprotein described by Meyer et al. (44), which is not homologous to ORF 17R.

The two exons of the highly conserved spliced gene (12L and 7L) can be used as an indicator of the evolutionary relatedness of herpesviruses. Costa et al. (12) have shown that these ORFs in HSV-1 (UL15) are transcribed late in infection to give a 2.7-kbp transcript after the removal of a 4-kbp intron. Davison and Scott (15) examined the conservation of nucleotide sequences at the splice donor-acceptor sites of these exons in VZV, EBV, and HSV-1. Figure 4 shows the sequences of HCMV (M. Chee and B. G. Barrell, manuscript in preparation) and HHV-6 added to this comparison. Table 5 shows measurements of the percentage similarity between the predicted amino acid sequences of the homologous portions of these ORFs in all possible pairwise comparisons between HSV-1, VZV, EBV, HVS, HCMV, and HHV-6. The proteins of the two alphaherpesviruses, HSV-1 and VZV, share 68% amino acid similarity, as do those of the gammaherpesviruses, EBV and HVS. HHV-6 shares 66% amino acid similarity with the betaherpesvirus HCMV, further evidence that at the molecular level this virus is more closely related to the betaherpesviruses than to the lymphotropic gammaherpesviruses.

DISCUSSION

The sequence presented here is the first large-scale analysis of sequences of the U1102 isolate of HHV-6. We have identified an overall colinearity between homologous genes of this virus and HCMV, a betaherpesvirus. Sequence comparisons by other workers have revealed both genetic colinearity between herpesviruses belonging to the same biological subgroup (13, 27) and large-scale rearrangements of blocks of conserved genes in herpesviruses belonging to different biological subgroups (16, 35). The region of HHV-6 presented in this report contains a number of highly conserved ORFs, including the major capsid protein and alkaline exonuclease. The conserved block of ORFs comprising 3L, 4L, and 5R is organized in the same orientation as the HCMV homolog and inverted compared with the EBV, HSV-1, and VZV homologs. The arrangement of the ORFs occurring within the intron of the spliced gene (on the opposite strand) is most similar to the arrangement seen in HCMV. Homologs of the HHV-6 ORFs 5R, 8R, and 14R are found only in HCMV. These observations, as well as the similarity of predicted amino acid sequences to those of HCMV, lead to the conclusion that HCMV and HHV-6 are as closely related as HSV-1 and VZV (13) or EBV and HVS (27). However, these genes of HHV-6 have a number of distinctive features; they are often smaller than their HCMV homologs, occur closer together, and overlap more often (Fig. 3), so that the smaller genome of HHV-6 (ca. 170 kbp [unpublished results]) appears to make more economical usage of DNA sequence than does the HCMV genome (230 kbp) in this region. In addition, two of the complete ORFs of the HHV-6 sequence (1L and 2L; Fig. 3) have no obvious counterparts in the genomes of any of the other human herpesviruses. It will be of interest to determine to what extent these sequences are conserved in other isolates of HHV-6.

Finally, some comments on the relationships between the molecular and biological properties of herpesviruses in general and HHV-6 in particular seem appropriate. The type species of the three biologically defined herpesvirus subgroups also comprise three clearly distinct groups in a molecular phylogenetic system (16, 27, 35). However, it is clearly possible to lose or gain nonhomologous functions, resulting in similar phenotypic properties. Thus, there is no necessary contradiction in a virus that may be a valid member of the biologically defined gammaherpesviruses having a betaherpesvirus as its closest relative in a molecular phylogeny. The major practical problems with the current biologically based classifications are the vagueness of the differentiating criteria and the absence of basic data on these biological properties, even when they are capable of being stated precisely. Since the only general theories of the relationships between genetic systems are essentially phylogenetic, a biologically based classification that is inconsistent with phylogenetic data is likely to be of limited usefulness. Recent studies of the genetic organization of the herpesvirus of turkeys and of Marek's disease virus provide an interesting illustration of the problems. These viruses have long been regarded as lymphotropic gammaherpesviruses because of the association of Marek's disease virus with lymphoproliferative disease, and much of the previous work interpreting their properties has proceeded by analogy with the association between EBV and B cells (47, 48). However, Buckmaster et al. (8) have shown that the genetic organizations of these two viruses are colinear with those of the alphaherpesviruses, HSV and VZV. Measurements of sequence similarity of the predicted protein products of these viruses also confirm a closer relationship to alphaherpesviruses than to sequenced gammaherpesviruses (8; unpublished results). These observations suggest that the lymphotropic properties of Marek's disease virus and herpesvirus of turkeys are unlikely to be determined by molecules homologous to those of EBV. The results we have presented here for HHV-6 provide a further example of a herpesvirus that has some lymphotropic properties but is more closely related to HCMV, a betaherpesvirus, than to other gammaherpesviruses. We suggest that comparisons between HHV-6 and HCMV are likely to prove more revealing of the basis for the divergence in their biological properties than are analogies with the properties of gammaherpesviruses such as EBV and HVS.

ACKNOWLEDGMENTS

We thank M. Jones for providing the plasmid pR9.1 sequenced in this study. G.L.L. thanks Carol Brown and Stacey Efstathiou for helpful discussion.

G.L.L. and M.C. thank the Commonwealth Scholarship Commission for financial support.

LITERATURE CITED

1. **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. Sequin, 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.
2. **Bankier, A. T., K. M. Weston, and B. G. Barrell.** 1987. Random cloning and sequencing by the M13/ dideoxynucleotide chain termination method. Methods Enzymol. **155**:51–93.
3. **Barton, G. J., and M. J. E. Sternberg.** 1987. Evaluation and improvements in the automatic alignment of protein sequences. Protein Eng. **1**:89–94.
4. **Barton, G. J., and M. J. E. Sternberg.** 1987. A strategy for the rapid multiple alignment of protein sequences. J. Mol. Biol. **198**:327–337.
5. **Biberfeld, P., B. Kramarsky, S. Z. Salahuddin, and R. C. Gallo.** 1987. Ultrastructure characterization of a new B-lymphotropic DNA virus (HBLV) isolated from patients with lymphoproliferative disease. JNCI **79**:933–941.
6. **Bird, A. P.** 1980. DNA methylation and the frequency of CpG in animal DNA. Nucleic Acids Res. **8**:1499–1504.
7. **Briggs, M., J. Fox, and R. S. Tedder.** 1988. Age prevalence of antibody to human herpesvirus 6. Lancet **i**:1058–1059.
8. **Buckmaster, A. E., S. E. Scott, M. J. Sanderson, M. E. G. Bournsnell, N. L. J. Ross, and M. M. Binns.** 1988. Gene sequence and mapping data from Marek's disease virus and herpesvirus of turkeys: implications for herpesvirus classification. J. Gen. Virol. **69**:2033–2042.
9. **Chappuis, B. B., K. Ellinger, F. Neipal, T. Kirchner, P. Kujath, B. Fleckenstein, and H. K. Muller-Hermelink.** 1989. Human herpesvirus 6 in lymphnodes. Lancet **i**:40–41.
10. **Chee, M., S.-A. Rudolph, B. Plachter, B. Barrell, and G. Jahn.** 1989. Identification of the major capsid protein gene of human cytomegalovirus. J. Virol. **63**:1345–1353.
11. **Chee, M. S., G. L. Lawrence, and B. G. Barrell.** 1989. Alpha-, beta- and gammaherpesviruses encode a putative phosphotransferase. J. Gen. Virol. **70**:1151–1160.
12. **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.
13. **Davison, A. J., and D. J. McGeoch.** 1986. Evolutionary comparisons of the S segments of the genomes of herpes simplex virus type 1 and varicella-zoster virus. J. Gen. Virol. **67**:597–611.
14. **Davison, A. J., and J. E. Scott.** 1986. DNA sequence of the major capsid protein gene of herpes simplex virus type 1. J. Gen. Virol. **67**:2279–2286.
15. **Davison, A. J., and J. E. Scott.** 1986. The complete DNA sequence of varicella zoster virus. J. Gen. Virol. **67**:1759–1816.
16. **Davison, A. J., and P. Taylor.** 1987. Genetic relations between varicella-zoster virus and Epstein-Barr virus. J. Gen. Virol. **68**:1067–1079.
17. **Dayhoff, M. O.** 1972. A model of evolutionary change in proteins, p. 89–99. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5. National Biomedical Research Foundation, Washington, D.C.
18. **Dayhoff, M. O.** 1972. Detecting distant relationships: computer methods and results, p. 101–110. *In* M. O. Dayhoff (ed.), Atlas of protein sequence and structure, vol. 5. National Biomedical Research Foundation, Washington, D.C.
19. **Deininger, P. L.** 1983. Random subcloning of sonicated DNA: application to shotgun DNA sequence analysis. Anal. Biochem. **129**:216–223.
20. **Downing, R. G., N. Sewankambo, D. Serwadda, R. Honess, D. Crawford, R. Jarrett, and B. E. Griffin.** 1987. Isolation of human lymphotropic herpesviruses from Uganda. Lancet **ii**:390.
21. **Draper, K. G., G. Devi-Rao, R. H. Costa, E. D. Blair, R. L. Thompson, and E. K. Wagner.** 1986. Characterization of the genes encoding herpes simplex type 1 and type 2 alkaline exonucleases and overlapping proteins. J. Virol. **57**:1023–1036.
22. **Efstathiou, S., U. A. Gompels, M. A. Craxton, R. W. Honess, and K. Ward.** 1988. DNA homology between a novel human herpesvirus (HHV-6) and human cytomegalovirus. Lancet **i**: 63–64.
23. **Eizuru, Y., T. Minamatsu, Y. Minamishima, M. Kikuchi, K. Yamanishi, M. Takahashi, and T. Kurata.** 1989. Human herpesvirus 6 in lymph nodes. Lancet **i**:40.
24. **Farrell, P. J.** 1987. Epstein-Barr virus (B95-8 strain), p. 99–107. *In* S. J. O'Brien (ed.), Genetic maps 1987, vol 4. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Fleckenstein, B., and R. C. Desrosiers.** 1982. Herpesvirus saimiri and herpesvirus ateles, p. 235–332. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
26. **George, D. G., W. C. Barker, and L. T. Hunt.** 1986. The Protein Identification Resource (PIR). Nucleic Acids Res. **14**:11–15.
27. **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.
28. **Honess, R. W.** 1984. Herpes simplex and 'the herpes complex': diverse observations and a unifying hypothesis. J. Gen. Virol. **65**:2077–2107.
29. **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 of differences in the states of their latent genomes. J. Gen. Virol. **70**:837–855.
30. **Honess, R. W., and D. H. Watson.** 1977. Unity and diversity in the herpesviruses. J. Gen. Virol. **37**:15–37.
31. **Josephs, S. F., A. Buchbinder, H. Z. Streicher, D. V. Ablashi, S. Z. Salahuddin, H.-G. Guo, F. Wong-Staal, J. Cossman, M. Raffeld, J. Sundeen, P. Levine, R. Biggar, G. R. F. Krueger, R. I. Fox, and R. C. Gallo.** 1988. Detection of human B-lymphotropic virus (human herpesvirus 6) sequences in B cell lymphoma tissues of three patients. Leukemia **2**:132–135.
32. **Josephs, S. F., S. Z. Salahuddin, D. V. Ablashi, F. Schacter, F. Wong-Staal, and R. C. Gallo.** 1986. Genomic analysis of human B-lymphotropic virus. Science **234**:601–603.
33. **Kishi, M., H. Harada, M. Takahashi, A. Tanaka, M. Hayashi, M. Nonoyama, S. F. Josephs, A. Buchbinder, D. V. Ablashi, F. Schacter, F. Wong-Staal, S. Z. Salahuddin, and R. C. Gallo.** 1988. A repeat sequence, GGGTTA, is shared by DNA of human herpesvirus 6 and Marek's disease virus. J. Virol. **62**:4824–4827.
34. **Knowles, W. A., and S. D. Gardner.** 1988. High prevalence of antibody to human herpesvirus-6 and seroconversion associated with rash in two infants. Lancet **ii**:912–913.
35. **Kouzarides, T., A. T. Bankier, S. C. Satchwell, K. Weston, P. Tomlinson, and B. G. Barrell.** 1987. Large-scale rearrangement of homologous regions in the genomes of HCMV and EBV. Virology **157**:397–413.
36. **Kozak, M.** 1984. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res. **12**:857–872.
37. **Lipman, D. J., and W. R. Pearson.** 1985. Rapid and sensitive protein similarity searches. Science **227**:1435–1441.
38. **Lopez, C., P. Pellett, J. Stewart, C. Goldsmith, K. Sanderlin, J. Black, D. Warfield, and P. Feorino.** 1988. Characteristics of human herpesvirus-6. J. Infect. Dis. **157**:1271–1273.
39. **Lusso, P., P. D. Markham, E. Tschachler, F. di Marzo Veronese, S. Z. Salahuddin, D. V. Ablashi, S. Pahwa, K. Krohn, and R. C. Gallo.** 1988. In vitro cellular tropism of human B-lymphotropic virus (human herpesvirus-6). J. Exp. Med. **167**:1659–1670.
40. **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 of the genome of herpes simplex virus type 1. J. Gen. Virol. **69**:1531–1574.
41. **McGeoch, D. J., A. Dolan, and M. C. Frame.** 1986. DNA sequence of the region of the genome of herpes simplex virus type 1 containing the exonuclease gene and neighbouring genes. Nucleic Acids Res. **14**:3435–3448.
42. **Mercer, J. A., C. A. Wiley, and D. H. Spector.** 1988. Pathogen-

esis of murine cytomegalovirus infection: identification of infected cells in the spleen during acute and latent infections. J. Virol. **62**:987–997.

43. **Messing, J., and J. Vieira.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
44. **Meyer, H., A. T. Bankier, M. P. Landini, C. M. Brown, B. G. Barrell, B. Ruger, and M. Mach.** 1988. Identification and prokaryotic expression of the gene coding for the highly immunogenic 28-kilodalton structural phosphoprotein (pp28) of human cytomegalovirus. J. Virol. **62**:2243–2250.
45. **Mizusawa, S., S. Nishimura, and F. Seela.** 1986. Improvement of the dideoxy chain termination method of DNA sequencing by use of deoxy-7-deazaguanosine triphosphate in place of dGTP. Nucleic Acids Res. **14**:1319–1324.
46. **Niederman, J. C., C.-R. Liu, M. H. Kaplan, and N. A. Brown.** 1988. Clinical and serological features of human herpesvirus-6 infection in three adults. Lancet **ii**:817–819.
47. **Nonoyama, M.** 1982. The molecular biology of Marek's disease herpesvirus, p. 333–346. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
48. **Payne, L. N.** 1982. Biology of Marek's disease virus and the herpesvirus of turkeys, p. 347–431. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
49. **Roizman, B.** 1982. The family Herpesviridae: general description, taxonomy, and classification, p. 1–23. *In* B. Roizman (ed.), The herpesviruses, vol. 1. Plenum Publishing Corp., New York.
50. **Salahuddin, S. Z., D. V. Ablashi, P. D. Markham, S. F. Josephs, S. Sturzenegger, M. Kaplan, G. Halligan, P. Biberfeld, F. Wong-Staal, B. Kramarsky, and R. C. Gallo.** 1986. Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. Science **234**:596–601.
51. **Sanger, F. S., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
52. **Saxinger, C., H. Polesky, N. Eby, S. Grufferman, R. Murphy, G. Tegtmeir, V. Parekh, S. Memon, and C. Hung.** 1988. Antibody reactivity with HBLV (HHV-6) in U.S. populations. J. Virol. Methods **21**:199–208.
53. **Senapathy, P.** 1988. Possible evolution of splice junction signals in eukaryotic genes from stop codons. Proc. Natl. Acad. Sci. USA **85**:1129–1133.
54. **Smith, R. F., and T. F. Smith.** 1989. Identification of new protein kinase-related genes in three herpesviruses, herpes simplex virus, varicella-zoster virus, and Epstein-Barr virus. J. Virol. **63**:450–455.
55. **Staden, R.** 1980. A new computer method for the storage and manipulation of DNA gel reading data. Nucleic Acids Res. **8**:3673–3694.
56. **Staden, R.** 1982. Automation of the computer handling of the gel reading data produced by the shotgun method of DNA sequencing. Nucleic Acids Res. **10**:4731–4751.
57. **Staden, R.** 1984. Measurements of the effects that coding for a protein has on a DNA sequence and their use for finding genes. Nucleic Acids Res. **12**:551–567.
58. **Staden, R.** 1986. The current status and portability of our sequencing handling software. Nucleic Acids Res. **14**:217–231.
59. **Tedder, R. S., M. Briggs, C. H. Cameron, R. Honess, D. Robertson, and H. Whittle.** 1987. A novel lymphotropic herpesvirus. Lancet **ii**:390–391.
60. **Weston, K., and B. G. Barrell.** 1986. Sequence of the short unique region, short repeats, and part of the long repeats of human cytomegalovirus. J. Mol. Biol. **192**:177–208.
61. **Yamanishi, K., T. Okuno, K. Shiraki, M. Takahashi, T. Kondo, Y. Asano, and T. Kurata.** 1988. Identification of human herpesvirus-6 as a causal agent for exanthum subitum. Lancet **i**: 1065–1067.