VIROLOGY 209, 29–51 (1995)

The DNA Sequence of Human Herpesvirus-6: Structure, Coding Content, and Genome Evolution

U. A. GOMPELS,*·†·‡·§·[1] J. NICHOLAS,§·** G. LAWRENCE,¶·†† M. JONES,‖ B. J. THOMSON,†·§ M. E. D. MARTIN,§ S. EFSTATHIOU,‡·¶ M. CRAXTON,§·¶ and H. A. MACAULAY*

*Viral Pathogenesis Unit, Department of Clinical Sciences, London School of Hygiene and Tropical Medicine, University of London, Keppel St, London WC1E 7HT, United Kingdom; Departments of †Medicine and ‡Pathology, University of Cambridge, Cambridge, United Kingdom; §National Institute for Medical Research, London, United Kingdom; ¶Laboratory of Molecular Biology, Cambridge, United Kingdom; ‖Royal Postgraduate Medical School, London, United Kingdom; **Department of Oncology, Johns Hopkins University, Baltimore, Maryland; and ††Macquarie University, Sydney, Australia*

Received November 14, 1994; accepted February 10, 1995

The complete DNA sequence was determined for strain U1102 of human herpesvirus-6, a CD4+ T-lymphotropic virus with disease associations in immunodeficient settings and a possible complicating factor in AIDS. The genome is 159,321 bp in size, has a base composition of 43% G + C, and contains 119 open reading frames. The overall structure is 143 kb bounded by 8 kb of direct repeats, DRL (left) and DRR (right), containing 0.35 kb of terminal and junctional arrays of human telomere-like simple repeats. Since eight open reading frames are duplicated in the repeats, six span repetitive elements and three are spliced, the genome is considered to contain 102 separate genes likely to encode protein. The genes are arranged colinearly with those in the genome of the previously sequenced betaherpesvirus, human cytomegalovirus, and has a distinct arrangement of conserved genes relative to the sequenced gammaherpesviruses, herpesvirus saimiri and Epstein–Barr virus, and the alphaherpesviruses, equine herpesvirus-1, varicella-zoster virus, and herpes simplex virus. Comparisons of predicted amino acid sequences allowed the functions of many human herpesvirus-6 encoded proteins to be assigned and showed the closest relationship in overall number and similarity to human cytomegalovirus products, with approximately 67% homologous proteins as compared to the 21% identified in all herpesviruses. The features of the conserved genes and their relative order suggested a general scheme for divergence among these herpesvirus lineages. In addition to the "core" conserved genes, the genome contains four distinct gene families which may be involved in immune evasion and persistence in immune cells: two have similarity to the "chemokine" chemotactic/proinflammatory family of cytokines, one to their peptide G-protein-coupled receptors, and a fourth to the immunoglobulin superfamily.

INTRODUCTION

Human herpesvirus-6 (HHV-6) is one of the newest characterized of the family of seven human herpesviruses. Like other herpesviruses it can establish a latent or persistent infection which remains for the lifetime of the host and can reactivate during immunosuppression. HHV-6 is one of the most widespread of the herpesviruses, infecting up to 90% of the population as infants where infection is asymptomatic or causes exanthem subitum, a mild skin rash (Yamanishi *et al.*, 1988; Okuno *et al.*, 1989; Ward *et al.*, 1993) with occasional severe or fatal complications (Asano *et al.*, 1990, 1992; Prezioso *et al.*, 1992; Drobyski *et al.*, 1994; Hall *et al.*, 1994; Ward and Gray, 1994). The virus has a cellular tropism for CD4+ T-lymphocytes (Tedder *et al.*, 1987; Takahashi *et al.*, 1989) and there is evidence for the monocyte/macrophage as a possible site for latency (Kondo *et al.*, 1991). HHV-6 strains were first isolated from blood samples from AIDS patients where the virus had reactivated (Salahuddin *et al.*, 1986; Downing *et al.*, 1987; Tedder *et al.*, 1987; Lopez *et al.*, 1988). Given the similar cellular tropisms of HIV and HHV-6, there has been speculation of interactions between these viruses in AIDS or the role of HHV-6 in immunodeficiencies, and some *in vitro* interactions have been observed (Ensoli *et al.*, 1989; Lusso *et al.*, 1989; Carrigan *et al.*, 1990; Levy *et al.*, 1990). In addition, limited *in vitro* replication in NK-cells and CD8+ T-lymphocytes has been shown to induce the HIV receptor, CD4, rendering the cells permissive for HIV replication (Lusso *et al.*, 1991, 1993). *In vivo* PCR analysis in HIV patients show disseminated HHV-6 infections and an inverse correlation with CD4-lymphocyte counts (Corbellino *et al.*, 1993; Fairfax *et al.*, 1994; Knox and Carrigan, 1994). Evidence for secondary, reactivated infections have also been observed in immunosuppressed bone marrow transplant patients and in these cases the virus has been associated with pneumonitis and bone marrow suppression (Carrigan *et al.*, 1991; Cone *et al.*, 1993; Drobyski *et al.*, 1993). Immunodeficiency with HHV-6-associated bone marrow suppression in a "normal" adult has been observed (Gompels *et al.*, 1994), which suggests a mechanism whereby HHV-6 reactivation may

[1] To whom correspondence and reprint requests should be addressed. Fax: +44 171 637 4314. E-mail: uag@hgmp.mrc.ac.uk.

0042-6822/95 $6.00

participate in AIDS. The pathogenesis of the virus is dependant on a balance between active lytic replication leading to cell death in permissive cells and the ability to remain in a latent state.

Human herpesviruses are classified in three subgroups, alpha, beta, and gamma, represented by herpes simplex virus (HSV), human cytomegalovirus (HCMV), and Epstein–Barr virus (EBV), respectively (Roizman *et al.*, 1992). In the virion the DNA is double stranded and linear, but during a latent infection of the host cell both HSV and EBV have virus genomes which appear "endless" (Efstathiou *et al.*, 1986); in EBV the genome has been shown to circularize. HHV-6 has characteristics of betaherpesviruses, sharing both encoded amino acid sequence similarities and overall gene organization with HCMV (Lawrence *et al.*, 1990; Neipel *et al.*, 1991; Gompels *et al.*, 1992; Nicholas and Martin, 1994). Studies of CpG suppression patterns (and matching TpG and CpA additions indicative of host-mediated methylation and repair/replication), showed lower frequencies only in the major immediate early gene regions in HCMV, murine CMV, rat CMV, and HHV-6 (Honess *et al.*, 1989; Martin *et al.*, 1991b; Sandford *et al.*, 1993). Thus, it has been proposed that the latent state of betaherpesviruses may involve control by localized methylation, which is distinct from models for alpha (HSV) or gamma (EBV) herpesviruses (Honess *et al.*, 1989; Martin *et al.*, 1991b). The genome of HHV-6 is smaller than that of HCMV, being 159 kb as compared with 230 kb, but comparisons show both are densely packed genomes with minimal splicing. HHV-6 has a distinct genetic structure composed of two G+C rich terminal direct repeats DRL (direct repeat left) and DRR (direct repeat right) which bound a long A+T rich unique sequence, UL (unique long) (Lindquester and Pellett, 1991; Martin *et al.*, 1991a). From initial sequencing studies it was shown that although both HHV-6 and HCMV share little nucleotide sequence similarity, they have a similar gene organization of "conserved" genes which encode homologous protein products (Lawrence *et al.*, 1990; Neipel *et al.*, 1991; Teo *et al.*, 1991; Efstathiou *et al.*, 1992; Gompels *et al.*, 1992; Nicholas and Martin, 1994). Both have an origin of lytic replication of the "complex type" in the same relative position upstream of the major DNA binding protein (Anders *et al.*, 1992; Masse *et al.*, 1992; Dewhurst *et al.*, 1993) and like the other herpesviruses, they appear to replicate by a rolling circle mechanism, generating head to tail concatemers which can be cleaved to unit length linear genomes as they are packaged into a preformed, proteinaceous "capsid" shell (Martin *et al.*, 1991a).

In this paper we report the complete DNA sequence of strain U1102 of HHV-6. This is one of the prototypic laboratory strains of HHV-6, first isolated from an Ugandan AIDS patient (Downing *et al.*, 1987). The sequence is examined and interpreted for its overall structure, coding content, and the genetic relationships between it and other herpesviruses. Most of the analyses for the UL region were first presented in August, 1992 at the XVII International Herpesvirus Workshop and parts of the sequence have been presented by us elsewhere (Lawrence *et al.*, 1990; Teo *et al.*, 1991; Efstathiou *et al.*, 1992; Gompels *et al.*, 1992; Thomson and Honess, 1992; Jones and Teo, 1993; Gompels and Macaulay, 1995; Lawrence *et al.*, 1995; Nicholas, 1994; Nicholas and Martin, 1994; Thomson *et al.*, 1994a).

MATERIALS AND METHODS

Cells and virus

The CD4+ T-lymphocyte cell line, JJhan, was grown in suspension culture at 37° in RPMI containing 2 m*M* glutamine, 100 U/ml penicillin, 100 μg/ml streptomycin, and 10% (vol/vol) fetal calf serum (myoclone, Gibco). Virus stocks were prepared by infection of JJhan cells with HHV-6 strain U1102 (Downing *et al.*, 1987) at a multiplicity of infection of 0.1 TCID50 and working stocks of virus used at passage 3. Phytohemaglutinen (PHA; 5 μg/ml)-stimulated cord blood lymphocytes were also used to propagate virus as described previously (Martin *et al.*, 1991a).

Virus, plasmid, cosmid, and bacteriophage DNA

The derivation of HHV-6 plasmid, cosmid, and bacteriophage clones covering the genome has been described (Martin *et al.*, 1991a). DNA preparations for sequence analysis were made using the alkaline–SDS method (Birnboim and Doly, 1979) of the following clones for subsequent sequence analyses: SMD2, SE6, SMD5, HD2, ED7, *Eco*J, *Bam*Q, ED4, HD1, HD9, *Xho*I, RPMS4.4, RPMS3.5, HD4.8, CB4, HD5.8, CB3, HD14, SMD6, ED4.5, SAD6, HD5, PstY, SAD3, SMD14, HD11, ED9, HD12, and CB11. The sequences of junctions between adjacent clones were determined either from overlapping independent clones or from polymerase chain reaction (PCR) amplification products using standard conditions and early passage virus DNA as template. *Sma*J, *Xho*K (from HD14), *Eco*L, and *Bam*J were derived from PCR amplified DNA. *Xho*K and *Sma*J were additionally cloned into pUC18 to make clones pSMJ6 and pXK.

DNA sequencing

DNA preparations were sonicated then gel purified to give random fragments averaging 600 bp in size. The ends were repaired with T4 DNA polymerase or *Escherichia coli* DNA polymerase I, Klenow fragment, and ligated with *Sma*I-digested and calf intestinal phosphatase-treated M13mp18 bacteriophage RF DNA using T4 DNA ligase. The ligated DNA was transfected into *E. coli* strains TG1 or JM109 and clones containing inserts were picked to prepare single-stranded DNA templates for sequencing following protocols based on the dideoxynucleotide chain termination method (Sanger *et al.*, 1977; Bankier *et al.*, 1988). *E. coli* DNA polymerase I, Klenow

fragment, was used for primer extensions with [^{35}S]dATP as the radioactive label (Biggin *et al.,* 1983). Reaction products were separated on 6% polyacrylamide, 8 *M* urea, Tris–borate-buffered gradient gels. Compressions were resolved either by incorporation of nucleotide analogues (deazaguanosine), using other polymerases (Sequenase or Taq polymerase), using internal primers, running gels at higher temperatures, or containing 10% formamide. Gels were dried onto Whatman 3MM chromatography paper and exposed to X-ray film.

DNA analyses

Autoradiographs were read using a digitizer, DNA Parrot (T&t Research, Ontario, Canada), or manually and the database was compiled using the SAP and DAP sequence assembly programs based on the Staden DB system (Staden, 1987), as adapted for UNIX, on SUN workstations (Gleeson and Staden, 1991). Each nucleotide in the sequence was read from an average of six overlapping random clones and 99.4% determined from both strands, with the total sequence derived from approximately 1×10^6 nucleotides of data. The sequence was analyzed using the Staden nucleotide and protein interpretation programs (nip, version 7.1, last update July 1993; and pip, version 5.0, February 1992) as adapted for UNIX (Gleeson and Staden, 1991) and the sequence analysis software of the GeneticsComputerGroup (GeneticsComputerGroup, 1991). The genome was compiled from overlapping sequence derived from plasmid, cosmid, and PCR clones using the assemble program of GCG and the compilation analyzed using the DNA interpretation and analysis program (DIANA, version 5 and 6/94, B. Barrell, Sanger Genome Centre, Cambridge, UK). Database comparisons were made using FASTA (version 1.6c24, optimise option, December, 1992) (Pearson and Lipman, 1988), BLAST (Altschul *et al.,* 1990), PROSRCH (Collins and Coulson, 1987) (version 1.1, 1990), and MPsrch (v 1.5; Storrock and Collins, 1993) programs and the Swissprot and EMBL databases (releases 28 and 39; 6/94).

RESULTS AND DISCUSSION

Genome organization

HHV-6 is the fifth human herpesvirus genome to be sequenced, those of the alphaherpesviruses, HSV-1 and varicella-zoster virus (VZV), gammaherpesvirus, EBV, and betaherpesvirus, HCMV, have been described previously (Baer *et al.,* 1984; Davison and Scott, 1986; McGeoch *et al.,* 1988; Chee *et al.,* 1990). It is only the second betaherpesvirus to be fully described, but has significantly different features from HCMV including a smaller, relatively A+T rich genome, with altered terminal repeat structure. It may be regarded as a subgroup 2 betaherpesvirus and similar subdivisions are seen in the alphaherpesviruses, VZV to HSV-1, and gammaherpesviruses, EBV, herpesvirus saimiri (HVS), equine herpesvirus 2 (EHV-2), and equine herpesvirus 5 (EHV-5) (proposed subgroup 2 and 3; Davison and Taylor, 1987; Albrecht *et al.,* 1992; Telford *et al.,* 1993). Preliminary sequence analyses on human herpesvirus 7 (HHV-7) (Berneman *et al.,* 1992) (J. Nicholas, unpublished results), the most recent member of the human herpesvirus family to be described, show a tight relationship with HHV-6 resembling that between HSV-1 and HSV-2 (McGeoch *et al.,* 1991; Roizman *et al.,* 1992). The more characterized HSV-1 is considered a prototype for the simplex-like viruses and similarly, HHV-6 may be considered the prototype for these newly described T-lymphotropic herpesviruses.

We have determined the sequence of the U1102 strain of HHV-6 (Downing *et al.,* 1987). Strain U1102 together with strains Z29 (Lopez *et al.,* 1988) and GS (Salahuddin *et al.,* 1986) are representative laboratory strains of HHV-6. Strain U1102 and GS are representative of variant A HHV-6 strains and Z29 of variant B strains (Ablashi *et al.,* 1993). Although, the nucleotide sequence differences between HHV-6 strains are less than that observed for HCMV and EBV strains, the HHV-6 strains can be arranged into at least two groups, variant A and B. However, further groupings have been observed within either variant class (Aubin *et al.,* 1991, 1993; Gompels *et al.,* 1993; Chou and Marousek, 1994). The complete linkage map and set of clones have been derived only for strain U1102 (Martin *et al.,* 1991a) and the virus can be grown in a number of CD4+ T lymphocyte cell lines. The complete sequence has been deposited with the EMBL and GenBank Data Libraries under Accession No. X83413.

The complete genomic sequence was determined after identification of the genomic termini present in the virion. This was done by comparisons of sequences derived from DRL terminal clones from purified virions of both strain U1102 and Z29, from concatemeric junctional clones (DRR–DRL) from infected cell DNA, and from clones containing junctions of UL to DRL and DRR (DRL–UL and DRR–UL) (Gompels and Macaulay, 1995; Thomson *et al.,* 1994a). The genome contains 159,321 bp, 8087 in the DRs (DRL, 1–8087 bp; DRR, 151,234–159,321 bp), and 143,147 bp in UL (from 8088 to 151,233 bp). The overall G+C content is 43% but is lower in UL (41%) and higher in DR (58%). This uneven base composition has been noted in the other sequenced herpesviruses, including equine herpesvirus 1 (EHV-1) and HVS (Baer *et al.,* 1984; Davison and Scott, 1986; McGeoch *et al.,* 1988; Chee *et al.,* 1990; Albrecht *et al.,* 1992; Telford *et al.,* 1992). HSV-1 and VZV contain an unpaired nucleotide at the 3′ end of each DNA strand (Mocarski and Roizman, 1982; Davison, 1984). This has not been identified here, since all sequence, including terminal clones, has been obtained from end-repaired fragments.

The HHV-6 genome does not show evidence for global CpG suppression as in the gammaherpesviruses and viral hosts (Honess *et al.,* 1989). In the vertebrate host, marked deficiency of CpG dinucleotide frequency is

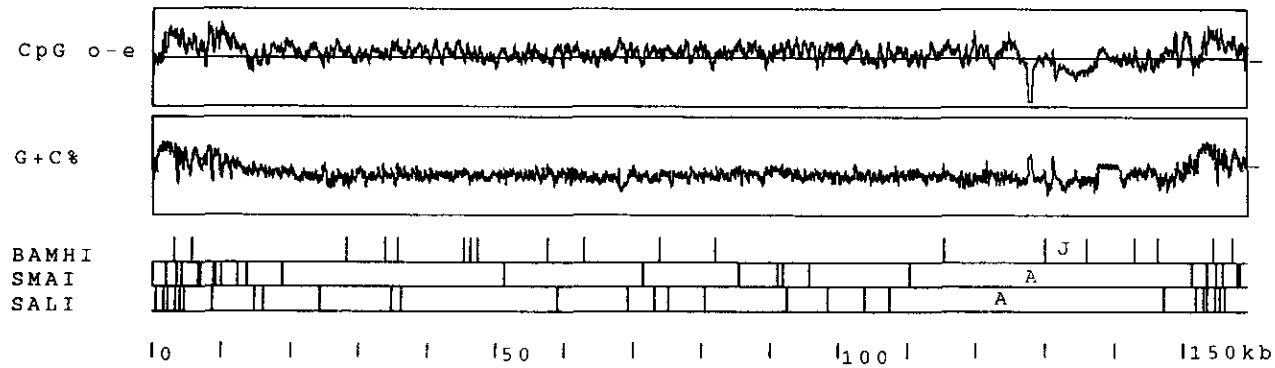


FIG. 1. HHV-6 genome has "local" rather than "global" suppression of CpG dinucleotide frequency. The top plot shows CpG frequencies as observed (o) − expected (e). The marked midline shows equal observed and expected frequencies and the scale is from a maximum of 25 to a minimum of −25. Span length is 401 and plot interval 20. The bottom plot shows overall G + C base composition. The marked midline shows 50% G + C with the scale maximum at 100% and minimum at 0%. Span length is 201 and plot interval 11. Restriction enzyme recognition sites are shown for *Bam*HI, *Sma*I, and *Sal*I and correspond to previous designations (Martin *et al.*, 1991a), although with improved resolution. *Bam*HI J, *Sma*I A, and *Sal*I A overlap the region of CpG suppression.

matched by an increase of TpG and CpA and is indicative of the mutagenic effects of C methylation at CG dinucleotides (Bird, 1980). Karlin and colleagues (Karlin *et al.*, 1994a) have noted that most viruses are CG suppressed, but in many cases this suppression seems unrelated to that described for vertebrates and certain herpesviruses, as corresponding increases in TG/CA frequencies were not observed. However, HHV-6, like HCMV and the other betaherpesviruses examined (SCMV and MCMV), have local CG suppression (Fig. 1) where CpG deficiency and matched TG/CA excess are only observed over the major immediate-early (IE) gene region. Presumably, this implied methylation functions here in gene control and latency in certain cell types or that the specific conformation of this region makes it accessible to methylases (Martin *et al.*, 1991b). It is possible that the similarly restricted pattern of CG suppression seen in HHV-6 and HCMV is related in some way to a biological function specific for betaherpesviruses, for instance, the site of latency of these viruses in cells of the monocyte/macrophage lineage, for which evidence has been presented (Kondo *et al.*, 1991; Taylor-Wiedeman *et al.*, 1991; Minton *et al.*, 1994; Taylor-Wiedeman *et al.*, 1994).

The complex direct repeats, DRL and DRR, contain simple repeats, unique sequences, and *cis*-acting regions involved in DNA packaging. Bounding each DR and thus the virus genome itself are 0.35-kb arrays of simple 6-mer repeats similar to human telomeric sequences (Kishi *et al.*, 1988; Martin *et al.*, 1991a). At the left end they are imperfect and of the pattern $[(TAACCC)_3(\text{related hexamer})_2]_4$ and then followed by related degenerate hexamers. On the right end there are 58 perfect tandem copies of this sequence (Gompels and Macaulay, 1995; Thomson *et al.*, 1994a). In strain Z29 the imperfect repeat follows the pattern $[(TAACCC)_3$ or $_4(\text{related hexamer})]_8$. The number of copies in either array may vary due to virus passage or plasmid cloning, and the imperfect array maps to the previously determined site of genomic heterogeneity, "het," at the left end of the direct repeats (terminal end of DRL and junction with UL of DRR) (Martin *et al.*, 1991a; Gompels and Macaulay, 1995; Thomson *et al.*, 1994a). These "het" repeats at the left end of the genome (DRL) are also the only set in the same orientation as the telomeric repeats found at the ends of human chromosomes. Thus, it is possible that human telomerase may add the human telomeric repeat, TTAGGG, at the 3′ end of HHV-6 DRL. This would lead to extension of single-stranded arrays and heterogeneity in DRL which may be copy/repaired to DRR (Gompels and Macaulay, 1995). However, the telomeric repeat arrays are not at the precise ends of the DRs; they are separated by 56 and 80 bp by *cis*-acting sequences, pac-1 and pac-2, involved in DNA packaging (Deiss *et al.*, 1986; Gompels and Macaulay, 1995; Thomson *et al.*, 1994a). It is debatable whether HHV-6 can act as template for human telomerase and if this plays any part in the heterogeneity observed or in latency. It has been noted that the repeats may play a role in segregation during cell division (Thomson *et al.*, 1994a) and that the left end repeats, although not precisely at the termini, may still be subject to telomerase repeat addition (Gompels and Macaulay, 1995) based on studies with tetrahymena telomerase (Harrington and Greider, 1991). This would be dependent on the cell type (some differentiated cells no longer synthesize telomerase) or whether the virus can induce telomerase. Interestingly, transformed, "immortalized" cell lines appear to induce telomerase (Greider, 1993). Although "het" regions are observed in virus which has been passaged in PBLs (Martin *et al.*, 1991a), "immortal" T-leukemic cell lines are often used to propagate HHV-6, and this may contribute to variability in this region.

The HHV-6 genome contains at least one origin for DNA replication within UL and possibly two others in DR. The origin of lytic replication, ori−lyt, in UL was first noted in both strains Z29 and U1102 by examination of se-

quence repeats upstream of the gene encoding major DNA binding protein, which is the position of the corresponding "ori" locus in HCMV (Hamzeh *et al.,* 1990; Anders *et al.,* 1992; Gompels *et al.,* 1992; Jones and Teo, 1992; Masse *et al.,* 1992; Dewhurst *et al.,* 1993). The region has now been functionally defined by transfection assays for replication of plasmids containing varying amounts of DNA from this locus (Dewhurst *et al.,* 1993). The minimal "ori" thus defined was 376 bp corresponding to position 67,617 to 67,993 bp in our U1102 sequence. The strains vary here by less than 5% at the nucleotide sequence level. The center of the "ori" is characterized by sequence with a relatively higher A+T content, repeated sequence including two 137-bp imperfect direct repeats (IDR, at positions 68,122 and 68,315 bp), HSV UL9 binding sites, and an adjacent GC-rich motif. It has been previously noted that HHV-6 encodes an origin of replication binding protein (OBP), with similar sequence and DNA binding specificity as that of HSV (UL9), but which is absent from HCMV (Gompels *et al.,* 1992; Inoue *et al.,* 1994; Lawrence *et al.,* 1995; Nicholas, 1994). Thus, the HHV-6 ori–lyt has features of both beta and alphaherpesvirus origins in having a larger set of repeat motifs (beta, "complex" ori) and sequences which bind the OBP (alpha ori). It has been speculated that another ori may be involved in plasmid maintenance with a role in latency (Thomson *et al.,* 1994a). Candidate oris may, in addition, be located in the center of each DR as there is an extremely AT rich region (82% A+T over 130 bp; centered at 3,589 in DRL and 154,812 in DRR) with adjacent multiple repeats, and this region is under evaluation (U. A. Gompels and H. A. Macaulay, unpublished results).

As well as the repetitive regions in DRL, DRR, and ori–lyt, there are three other clusters of major reiterations. These are located at the right end of UL and include the previously described SR, TG, and Kpn repeats at positions 127, 131, and 138 kb in the genome, respectively (Martin *et al.,* 1991a,b; Thomson and Honess, 1992; Nicholas, 1994). Here they are designated R1, R2, and R3, but the precise boundaries are not clear since there are degenerate repeats which confound their delineation. R1, the SR repeats, are reiterations which encode an SR domain (Birney *et al.,* 1993; Nicholas, 1994) in the HCMV IE2 homologue in HHV-6 (U86) and they extend between 127,497 and 128,180 bp. R2 are composed of simple TG dinucleotide repeats between 131,021 and 132,253 bp, where T+G is 78% of the base composition in this region, creating large reading frames open in all six frames. This unusual composition, may have contributed to the inability to clone this region in standard vectors and the sequence was only derived from PCR amplifications (Martin *et al.,* 1991a,b). R3 is composed of the 110-bp tandem kpn containing repeats. Kpn site containing repeats extend from 137,993 to 140,896 bp in the sequence presented here.

Tandem telomere-like repeats, TAACCC, are not found outside DRL or DRR. However, within UL, monomeric copies can be identified. These are distributed evenly throughout UL, but in a complementary fashion either side of the UL ori with TAACCC before and GGGTTA after the ori (Gompels and Macaulay, 1995). A model to account for this is that an RNA molecule with the TAACCC repeat acts in replication of the lagging strand in synthesis of Okazaki fragments. Over time, repair replication would generate a complementary compositional bias either side of the ori and this is indeed observed in UL (Gompels and Macaulay, 1995). In retroviruses, host tRNAs are packaged and used as RNA primers which have complementarity to the retrovirus genome (Suzuki *et al.,* 1993). Intriguingly, tetrahymena telomerase has an RNA component with complementarity to the telomeric repeat which is used as template for telomerase in repeat addition (Greider, 1993). A model which includes such an RNA molecule with HHV-6 repeats (UAACCC) would imply that repeat replication in the DRs (at het) and lagging strand synthesis may have similar features in HHV-6 (Gompels and Macaulay, 1995).

Gene arrangement

Potential genes encoded by HHV-6 strain U1102 were identified by several criteria: open reading frames (ORF) with initiating methionine codons (ATG) greater than 80 bp and nonoverlapping by 75%, probability of coding based on positional base preference methods using the compositional bias (Gleeson and Staden, 1991), similarity of encoded amino acid sequences to other herpesvirus or cellular proteins, and transcriptional initiation and termination signals. For certain ORFs significantly larger than 80 bp, where no initiation codons were identified within the first 30% of the ORF, identification of splice donor and acceptor sites were used. One hundred nineteen ORFs were identified (Fig. 2, Table 1). In the DRs they are designated DR1–DR8 and in UL they are U1–U100. Three ORFs span the telomeric repeats. These are at the DRL terminus, the DRL/UL junction, and the DRR/UL junction and are designated LT1, LJ1, and RJ1. A tentative designation is U88 in R2. This region is open in all six frames, but only U88 is designated as it is the largest ORF with an initiating methionine. In the absence of additional splicing data, ORFs designated are the largest which fit the above overall criteria. DR has a higher G+C composition (58%) than UL (41%) and thus most of the DR ORFs, particularly DR2, encode proteins with a bias to G+C rich codons, for example proline. Therefore, similarity searches with DR-encoded sequences identify matches with other proline rich genes in EBV (EBNAs) and HSV (IE genes) which may be spurious.

The arrangement of 119 ORFs is shown in Fig. 2. Of these, 102 seem likely to be distinct genes. DR1–8 are duplicated in DRL and DRR, ORFs U60 and U66 are homologous to the two exons of HSV UL15, the late spliced gene (Costa *et al.,* 1984; Dolan *et al.,* 1991), U17 appears to be an exon spliced to U16 (Nicholas and Martin, 1994),

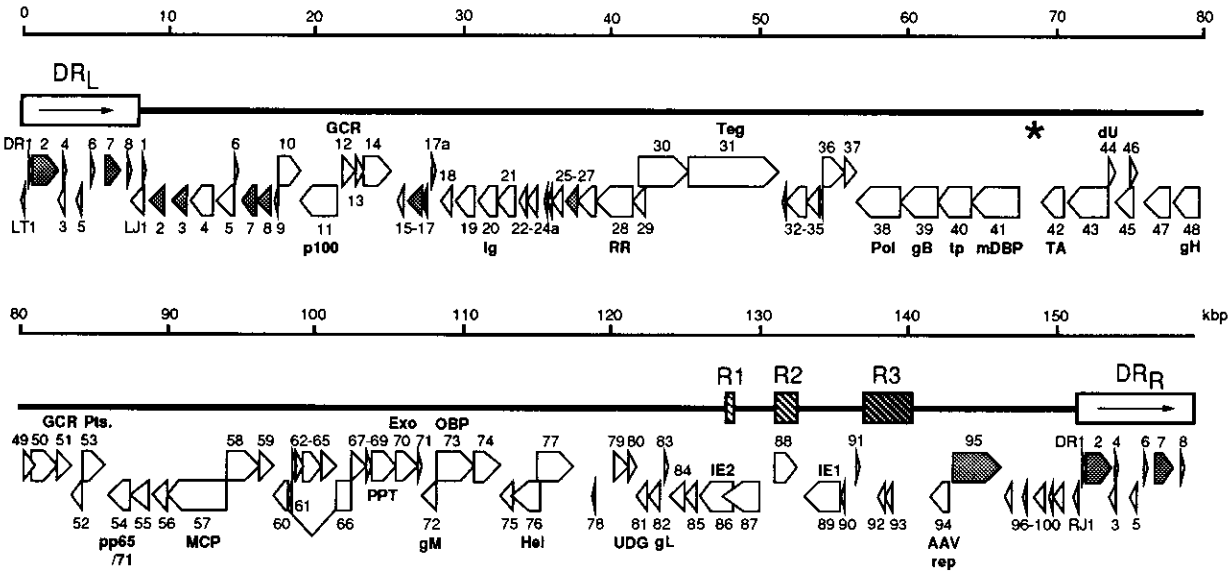


FIG. 2. Predicted HHV-6 gene organization. Repeat regions are boxed, DRL, DRR, R1, R2, R3, and the long unique (UL) region is indicated by a solid line. Protein coding regions are indicated as open arrows and are numbered DR1–DR8 in the direct repeats and U1–U100 in UL. The ori–lyt is indicated by a star on UL. Abbreviations are: GCR (G-protein coupled receptor), RR1 (large subunit of ribonucleotide reductase), TEG (tegument protein), POL (DNA polymerase), tp (transport protein), mDBP (major single-stranded DNA binding protein), TA (conserved herpesvirus transactivator), Pts (protease/assembly protein), exo (alkaline exonuclease), OBP (origin binding protein), UDG (uracil DNA glycosylase), Hel (helicase), MCP (major capsid protein), AAV rep (Adeno-associated virus replication protein homologue), IG (immunoglobulin superfamily). These are further described in Table 1 and the text. The US22 gene family are shaded and are listed in Table 2.

and U90 is an exon spliced to U89 with other sequences (Martin *et al.*, 1991b; Nicholas, 1994; Schiewe *et al.*, 1994). Therefore, the total number of distinct ORFs is 108 and these are listed in Fig. 3. If those spanning the telomeric and R1,2,3 repeats are eliminated, 102 genes are likely to encode protein. Recent analysis of spliced IE genes at the right end of UL by cDNA cloning and sequencing have identified another spliced gene, U91, and it is expected that other spliced genes may be identified in this region (Schiewe *et al.*, 1994). HHV-6 U53 is related to HCMV UL80 and HSV UL26 which are expressed as a full-length product and as an internally initiated 3′ coterminal mRNA. This transcript is in-frame, encoding truncated products, UL80a and UL26.5, respectively. This pattern of expression is also likely in HHV-6 as U53 has an internal initiating methionine located in a similar position; thus, the smaller overlapping gene has been designated U53a (Table 1). Currently, little is known about transcriptional organization except for the spliced genes described above located on either end of UL and in DRR; more genes may therefore be characterized with further analyses of these spliced regions. With the exception of the late conserved spliced gene, there appears to be little evidence for splicing in the central core of the genome and overall it is densely packed (Fig. 2).

Betaherpesvirus gene arrangement

The overall arrangement of genes is similar to that of HCMV, but diverges at either end (Fig. 3). In UL similarity to HCMV genes begins at U2 and terminates at U86 (9,816 to 125,980 bp), corresponding to HCMV UL23 to UL122 (27,866 to 170,878 bp). Thus, within UL the betaherpesvirus gene arrangement spans 116,175 bp in HHV-6 and 143,012 bp in HCMV. This means that the Human Leukocyte Antigen (HLA) class I homologue, UL19, and most of the glycoprotein gene families, RL11, US6, and US12, present at the left end of the HCMV genome and at the right end within US, are lacking in HHV-6. The majority of the HHV-6 specific genes are located outside this betaherpesvirus gene arrangement but there are also some within (Fig. 3, Table 1). Many are designated glycoproteins such as U20, U21, U22, U23, U24, and U73 which are in similar positions to glycoprotein genes or exons in HCMV but without detectable similarity. The common "core" of genes conserved in all herpesviruses, determined by similar encoded amino acid sequence and described in the following sections, extends from U27 at 37,800 bp to U81 at 122,577 bp. Although the identification of common herpesvirus "core" genes is limited by methods used to detect similarity, the analyses shown here gives HHV-6 a core of 86 kb, the most compact of all the herpesviruses examined (Figs. 3 and 4; Table 3).

The DR regions of HHV-6 appear to be related to the short (S) region of HCMV which includes IRS, US, and TRS (Fig. 2). Although the telomeric repeats in HHV-6 DR are absent from the S region in HCMV, both DR and S share certain features. Both are located at the ends of the genome, contain the pac1 and pac2 sites for DNA packaging, and encode members of the HCMV US22 gene family. This gene family extends into UL (Fig. 2; Table 2) and includes genes which encode a "CC" containing motif (Kouzarides *et al.*, 1987; Chee *et al.*, 1990;

Efstathiou *et al.*, 1992; Thomson and Honess, 1992). The members in DR, DR2, and DR7 are most closely related to US26 and US22, respectively; therefore, the DR resembles the inverse orientation of US of HCMV. U95 is the positional homologue of the US22 member in the repeats bounding HCMV US, TRS, and IRS. Such genetic expansion/contractions of genes within and adjacent to terminal repeat regions, which make unique genes appear "captured" by repeated regions, have been described for the US regions of the alphaherpesviruses (Davison and McGeoch, 1986; McGeoch, 1990; Telford *et al.*, 1992).

The presence of gene families may be a general feature of betaherpesviruses. HCMV is noted for its extensive gene families (Chee *et al.*, 1990). HHV-6, also has a number of gene families, some of which are related to those in HCMV. These families include the US22, DR1/6, U4/5, G-protein coupled receptor (GCR), and immunoglobulin gene families (Ig). The US22 family is the most extensive and has 11 members (two repeated in DR) and are related to varying extents to the 12 members in HCMV (Table 2) (Kouzarides *et al.*, 1988; Chee *et al.*, 1990). In addition, one has been identified in murine cytomegalovirus, MCMV IE2 (Messerle *et al.*, 1991). Interestingly, although there are positional homologues to members of this family (marked P on Table 2), the closest similarity is not always to the positional homologue but often to another member of the family. This suggests conservation of domains with similar functions rather than the genes themselves or possible interactions between members of the family. Some members of this group are spliced and expressed as immediate early proteins, U16/17, HCMV UL36, and MCMV IE2, but it is not known if this extends to all of the family (Kouzarides *et al.*, 1988; Messerle *et al.*, 1991; Nicholas and Martin, 1994). There is evidence for a role in transcriptional activation for U16 (Geng *et al.*, 1992; Nicholas and Martin, 1994), DR7 (Thompson *et al.*, 1994), HCMV UL36 (Colberg-Poley *et al.*, 1992), and HCMV IRS1/TRS1 (Stasiak and Mocarski, 1992), but again it is not known if all US22 gene products are involved in transactivation. At least one member, DR7, appears to have a role in cellular transformation *in vitro* (Thompson *et al.*, 1994).

Other gene families in HHV-6 and HCMV appear to have arisen as gene duplications which have diverged (Tables 1 and 3). DR1 and DR6 both share a CXC motif with a third cysteine aligning. Both are also adjacent to US22 family genes which indicates some symmetry within the DRs. U4 and U5 are tandem duplications and positional homologues of HCMV UL27. Conversely, some gene duplications observed in HCMV are absent in HHV-6. For example, U14 is related to both HCMV UL25 and UL35. The positional homologue to U14 in HCMV is UL35, but it is more closely related to HCMV UL25, suggesting that this gene may be the original from which UL35 is derived. Similarly, U54 is related to the tandem duplication in HCMV of UL82 and UL83, but further comparisons using the MPsrch and Prosrch programs show that it is more closely related to UL83.

Two further HHV-6 gene duplications have both HCMV and cellular homologues. U12 and U51 are related to all GCR and are distantly related to GCR homologues in HCMV, UL33, US22, and US28, as well as UL78, which was a previously undetected member of this family (U. A. Gompels *et al.*, in preparation). U20 and U85 are both glycoproteins with domains similar to the immunoglobulin superfamily. The GCR and Ig family proteins are described further in the next section.

Gene function

The coordinates of the HHV-6 ORFs and general characteristics of the products of primary translations are listed in Table 1. Where there is no additional data on transcription it is assumed that translation starts at the first ATG in each ORF. Where new splice donor/acceptor sites were identified the whole ORF is listed. Homologues of HHV-6 proteins in the other sequenced herpesviruses, EHV-1, HSV-1, HVS, EBV, and HCMV, were identified using the FASTA program (Pearson and Lipman, 1988) with the optimize option (ordering by optimized scores) and a word size of 2 (in some cases 1 was used in addition). These are also listed (Table 3). Positional homologues are also indicated by a P where genes are in a similar location encoding products with similar overall character, size, and function (where known), but which have undetectable amino acid sequence similarity using these methods. FASTA scores greater than 100 were considered to indicate significant homology. The HHV-6 genes have a colinear arrangement with their counterparts in HCMV as well as encoding products with closer overall similarity in comparison with the other sequenced herpesvirus genomes. This protein sequence similarity underlines the relationship between these viruses as revealed by comparison of genome organization discussed above. The properties of HHV-6 proteins are derived in a limited number of cases from direct experimental data or extrapolated from functions defined for HCMV or in the more distantly related HSV.

Membrane proteins

The primary translation products of eight HHV-6 genes share features characteristic of class I membrane glycoproteins (Von Heijne and Gavel, 1988). These include a hydrophobic signal sequence near the amino terminus for translation of the mRNA on membrane-bound ribosomes, sites for N-linked glycosylation (NXT/S), and a carboxy-terminal hydrophobic sequence followed by basic residues which can act as a transmembrane anchor. U39 and U48 are related to gB and gH, respectively, which have been described previously and are members of the set of glycoproteins conserved in all herpesviruses. Both have roles in virus infection and cellular spread similar to that shown for other herpesvirus homologues.

TABLE 1

Feature Table of HHV-6 Genes and Encoded Proteins[a,b,c]

name	strand	start	stop	aa:mw	old name	closest homologue	gene family	gene block	- comment
LT1	C	338	3	112:12816					
DR1	-	501	791	97:11677			DR1/DR6		- CXC motif
DR2	-	791	2650	620:67151		HCMVUS26*	HCMVUS22		
DR3	C	2979	2404	192:19404					
DR4	-	2746	3045	100:10562					
DR5	C	4171	3737	145:15637					
DR6	-	4725	5033	103:12129			DR1/DR6		- CXC motif
DR7	-	5629	6717	363:40831		HCMVUS22*	HCMVUS22		- transformation, transactivator
DR8	-	7237	7566	110:12766	SJRF1				- SR domain
LJ1	C	8432	7470	321:36736	SJLF1				- across DRL junctional telomeric repeats
U1	-	8245	8613	123:14338	SJRF2				- SR domain
U2	C	9816	8719	366:39815	SHL1	HCMVUL23*	HCMVUS22		
U3	C	11276	10158	373:43700	SHL2	HCMVUL24*	HCMVUS22		
U4	C	13092	11488	535:61963	SHL3	HCMVUL27*	U4/U5		
U5	C	14548	13217	444:51587	SSL1	HCMVUL27*	U4/U5		
U6	-	14619	14864	82:8878					
U7	C	15936	14911	342:39357	SSL2	HCMVUL28*	HCMVUS22		
U8	C	17091	16024	356:41103	SFL1	HCMVUL29*	HCMVUS22		
U9	C	17552	17241	104:11819	SFL2				
U10	-	17604	18911	436:50123	SFR1,P1RF0	HCMVUL31*			
U11	C	21578	18969	870:97073	P1LF1	HCMVUL32*			- pp100 major antigenic structural protein, basic phosphoprotein, BPP
U12EX	-	21680	21710	347:39747+					- U12 exon,spliced donor/acceptor 21710,21800+
U12	-	21856	22809	318:36611	P1RF1	HCMVUL33*	GCR		- G-protein coupled receptor homology;EBV induced EBI1
U13	-	22898	23215	106:12003	EFRF1				
U14	-	23316	25142	609:69561	EFRF2	HCMVUL25*			- homology to HCMV'UL25/35' family
U15	C	25992	25663	110:13416	EFLF3				
U16EX	C	27349	26262	312:36092+		HCMVUL36EX2*			- IE-B;U16exon,spliced acceptor/donor 27034,27187
U16	C	-27116	26262	285:32908	EFLF2,IE-B	HCMVUL36*	HCMVUS22		- IE-B;transactivator,
U17	C	27349	26951	133:15538	EFLF1	HCMVUL36EX1*	HCMVUS22		- IE-B;U16exon1?,acceptor/donor 27034,27187
U18	C	29389	28511	293:33219	EJLF6	HCMVUL37EX3*			- IE-B;homology to HCMV IE glycoprotein
U19	C	30818	29652	389:43317	EJLF4	HCMVUL38*			- IE-B
U20	C	32337	31072	422:48689	EJLF3		IG		- glycoprotein;Ig chain C domain
U21	C	33641	32343	433:49396	EJLF2				- glycoprotein
U22	C	34347	33742	202:23409	EJLF1				- glycoprotein
U23	C	35085	34378	236:26524	EqLF1				- glycoprotein
U24	C	35655	35395	87:10092	EoLF1				- glycoprotein exon? glyco site,TM
U25	C	36814	35867	316:37097	EPLF3	HCMVUL43*	HCMVUS22		
U26	C	37809	36925	295:33336	EPLF2				
U27	C	38978	37800	393:44811	EPLF1	HCMVUL44***		I	- pp41,pol processivity,transactivator,HCMV ICP36
U28	C	41434	39023	804:93354	P2LF2	HCMVUL45***		I	- large sub-unit ribonucleotide reductase;RR1
U29	C	42356	41460	299:34205	P2LF1	HCMVUL46***		I	- capsid assembly and DNA maturation; minor capsid protein, mCP
U30	-	41884	45129	1082:124060	P2RF1	HCMVUL47***		I	- HCMV capsid assembly,myosin
U31	-	45150	51380	2077:239950	HHRF1	HCMVUL48***		I	- large tegument protein,high molecular weight protein, HMWP
U32	C	51721	51458	88:9742					
U33	C	53135	51726	470:54701	XILF3	HCMVUL49***		I	- capsid protein
U34	C	53916	53089	276:31672	XILF2	HCMVUL50***		I	- possible virion protein,TM
U35	C	54253	53936	106:12441	XILF1	HCMVUL51* *			
U36	-	54252	55703	484:56409	XIRF1	HCMVUL52***		I	- probable virion protein
U37	-	55710	56501	264:30846	XIRF2	HCMVUL53***		I	- nuclear phosphoprotein?
U38	C	59588	56553	1012:115822	XILF0,Pol	HCMVUL54***		II	- DNA polymerase
U39	C	62080	59591	830:93240	gpB	HCMVUL55***		II	- glycoprotein B
U40	C	64214	62037	726:82879	tp	HCMVUL56***		II	- transport protein
U41	C	67620	64225	1132:127765	mdbp	HCMVUL57***		II	- major DNA binding protein
U42	C	70598	69057	514:59771		HCMVUL69***		III	- conserved herpesvirus transactivator-
U43	C	73405	70826	860:100082		HCMVUL70***		III	- helicase/primase complex; HSV primase
U44	-	73446	74084	213:24297		HCMVUL71***		III	
U45	C	75218	74091	376:43399	BHLF3	HCMVUL72***		III	- putative dUTPase
U46	-	75291	75542	84:9779	BHRF1	HCMVUL73***		III	- membrane/secreted protein
U47	C	77867	75915	651:73251	variable gp	HCMVUL74*			- membrane/secreted glycoprotein
U48	C	80118	78037	694:79544	gH,BHLF1	HCMVUL75***		IV	- glycoprotein H
U49	-	80277	81032	252:29293	BHRF2	HCMVUL76***		IV	- fusion protein
U50	-	80812	82476	555:63596	BHRF3	HCMVUL77***		IV	- virion protein
U51	-	82574	83476	301:34715	XKRF1	opioid^R/HCMVUL78*	GCR		- G-protein coupled receptor homology; HVS GCR
U52	C	84274	83501	258:30031	XKLF1	HCMVUL79*			
U53	-	84281	85864	528:58637	0R,XKRF1	HCMVUL80***		IV	- protease; U53a in-frame assembly protein, AP
U54	C	87427	86054	458:51499	1L	HCMVUL82/83*			- tegument pp65/72K, possible transactivator IE genes
U55	C	88803	87508	432:50172	2L	HCMVUL84*			
U56	C	89873	88986	296:33465	3L	HCMVUL85***		V	- probable capsid protein
U57	C	93912	89878	1345:151953	4L,mcp	HCMVUL86***		V	- major capsid protein,MCP

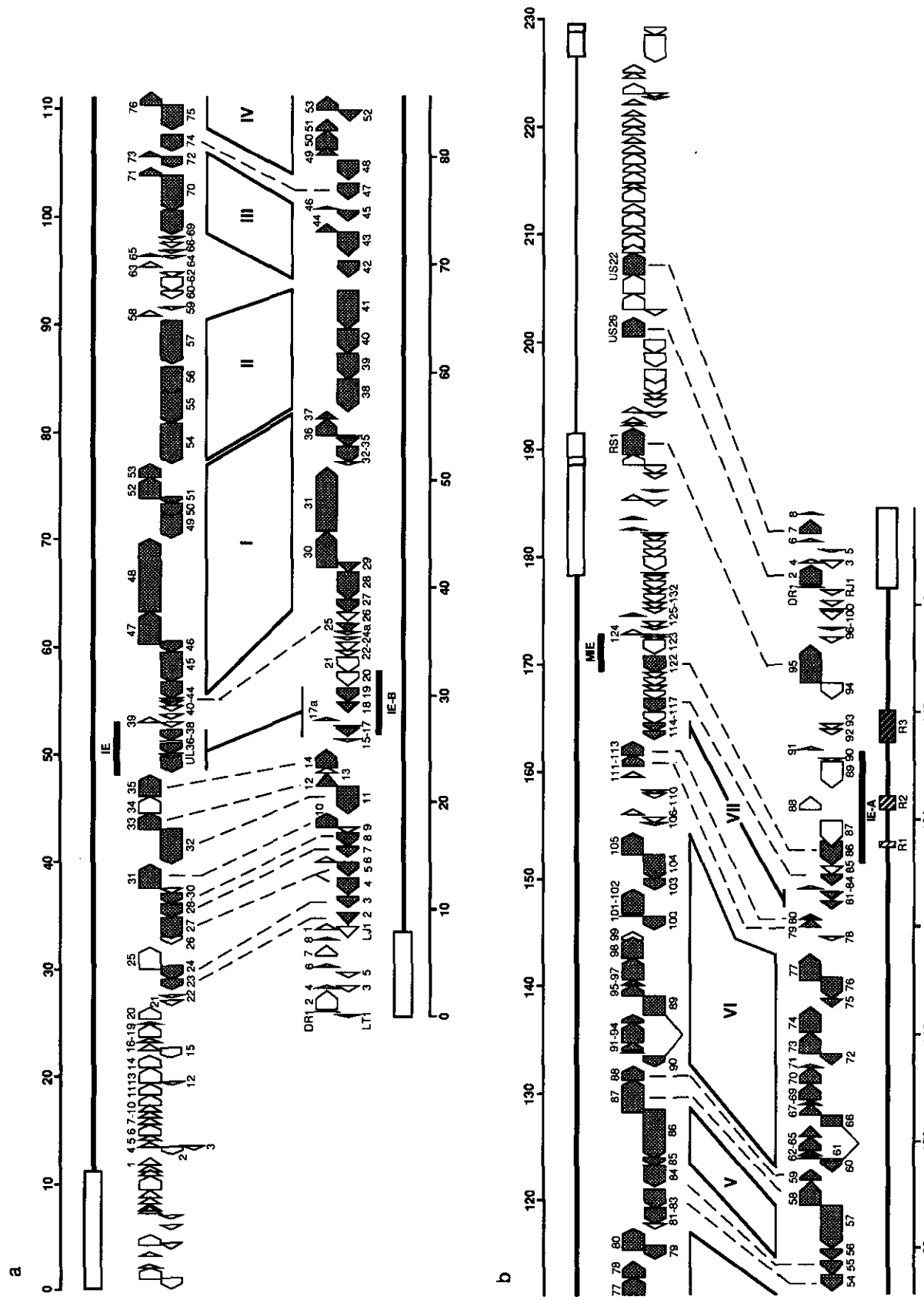
TABLE 1 — *Continued*

name	strand	start	stop	aa:mw	old name	closest homologue	gene family	gene block	- comment
U58	-	93924	96239	772:88748	5R	HCMVUL87**			
U59	-	96239	97288	350:39884	6R	HCMVUL88*			
U60	C	98256	97291	322:36084	7L	HCMVUL89EX2***		VI	- late spliced gene(U60/66),possible DNA packaging protein
U61	C	98578	98234	115:13578					
U62	-	98427	98681	85:9579	8R	HCMVUL91*			
U63	-	98632	99279	216:24783	9R	HCMVUL92**			
U64	-	99260	100585	442:51392	10R	HCMVUL93***		VI	
U65	-	100545	101549	335:37878	11R	HCMVUL94***		VI	
U66	C	102486	101572	305:35930	12L	HCMVUL89EX1***		VI	- late spliced gene(U60/66),possible DNA packaging protein
U67	-	102458	103516	353:39533	13R	HCMVUL95***		VI	
U68	-	103519	103860	114:13075	14R	HCMVUL96**			
U69	-	103866	105551	562:63718	15R	HCMVUL97***		VI	- ganciclovir kinase; conserved phosphotransferase
U70	-	105562	107025	488:56646	16R	HCMVUL98***		VI	- alkaline exonuclease
U71	-	106965	107195	77:8469	17R				- position HCMV pp28k,HSV myristilated virion protein
U72	C	108312	107281	344:38993	18L	HCMVUL100***		VI	- *integral membrane protein, gM*
U73	-	108325	110664	780:89719	19R,HDRF0	HSVUL9+			- origin binding protein
U74	-	110636	112621	662:76318	HDRF1	HCMVUL102***		VI	- helicase/primase complex
U75	C	113408	112662	249:28762	HDLF2	HCMVUL103***		VI	
U76	C	115305	113320	662:77236	HDLF1	HCMVUL104***		VI	- possible virion protein
U77	-	115100	117571	824:93288	HDRF2	HCMVUL105***		VI	- helicase/primase complex; helicase
U78	C	119038	118712	109:12725	EDLF5				
U79	-	120164	121195	344:39273	EDRF1	HCMVUL112*			- HCMV in vitro replication;spliced
U80	-	-121170	121763	198:22256	EDRF2	HCMVUL113*			- HCMV in vitro replication; spliced
U81	C	122577	121813	255:29039	EDLF4	HCMVUL114***		VII	- uracil-DNA glycosylase
U82	C	123405	122656	250:28962	gL,EDLF3	HCMVUL115***		VII	- glycoprotein gL; gH accessory protein
U83	-	123528	123818	97:10411	EDRF3				- CC chemokine?
U84	C	124953	123928	342:39557	EDLF2	HCMVUL117*			- spliced in HCMV
U85	C	125853	124984	290:32901	EDLF1		IG		- OX-2 homology; glycoprotein
U86	C	-128136	125992	715:80040	BCLF1	HCMVUL122*			- IE-A;HCMVIE2 homology, SR domain repeats
U87	C	130043	127554	830:91391	BCLF0				- IE-A;glycoprotein?, highly charged,pro repeats
U88	-	131034	132272	413:44136					- IE-A;open all frames, cys repeats
U89	C	135610	133094	839:93712	pRF3/4;RF2				- IE-A;HCMV IE1 position, transactivator
U90	C	135948	135667	94:10651	pRF2;RF1				- IE-A; spliced U89
U91	-	-136485	136826	114:12897					- IE-A;spliced antisense IE1?
U92	C	-138492	138052	147:15878					- kpn repeats, part duplicate U93
U93	C	-139124	138534	197:21371					- kpn repeats, part duplicate U92
U94	C	142866	141397	490:55849	HCLF2	AAV2 Rep 68/78			- parvovirus replication, transactivation
U95	-	142941	146303	1121:124069	HCRF2	MCMVIE2*	HCMVUS22		- positional/sequence homolgue MCMVIE2
U96	C	146940	146644	99:12128	HCLF1				
U97	C	148077	147811	89:10384					-
U98	C	-149391	148744	296:24698					
U99	C	149766	149488	93:10559					- signal sequence
U100	C	150437	149871	189:21609	gp82/105				- spliced glycoprotein gp82/105
RJ1	C	151571	151143	143:16221					- *across DRR junctional telomeric repeats*
DR1	-	151734	152024	97:11677			DR1/DR6		- CXC motif
DR2	-	152024	153883	620:67151		HCMVUS26*	HCMVUS22		
DR3	C	154212	153637	192:19404					
DR4	-	153979	154278	100:10562					
DR5	C	155404	154970	145:15637					
DR6	-	155958	156266	103:12129			DR1/DR6		- CXC motif
DR7	-	156862	157950	363:40831		HCMVUS22*	HCMVUS22		- transformation, transactivator
DR8	-	158470	158799	110:12766					- SR domain

[a] Start is the first base in ATG (or its complement), except as indicated by a dash, where first base in ORF or exon is listed. Stop is third base in stop codon. + indicates spliced gene encoded molecular weight.

[b] ***conserved in alpha, beta, and gammaherpesviruses; **conserved in beta and gammaherpesviruses; *conserved in betaherpesviruses; * *conserved in beta and alphaherpesviruses.

[c] Properties for proteins are derived from homologues described in HSV-1, VZH, EHV-1, or HCMV (Davison and Scott, 1986; McGeoch *et al.*, 1988; Chee *et al.*, 1990; Telford *et al.*, 1992) with additional data (not exhaustive, also see text) for the following HCMV genes UL33 (Neote *et al.*, 1993), UL36 (Colberg-Poley *et al.*, 1992), UL44, UL54, UL57, UL70, UL84, UL102, UL105, UL112, UL113 (Pari and Anders, 1993), UL115 (Kaye *et al.*, 1992, UL69 (Winkler *et al.*, 1994), UL82 (Liu and Stinski, 1992), UL97 (Littler *et al.*, 1992; Sullivan *et al.*, 1992), UL80 (Jones *et al.*, 1994); for MCMV gene IE2 (Koszinowski *et al.*, 1986; Messerle *et al.*, 1991) for HSV genes UL10 (Baines and Roizman, 1993), UL15 (Poon and Roizman, 1993), UL18, UL19, UL35, UL38, UL26, UL26.5 (Liu and Roizman, 1991; Newcomb *et al.*, 1994; Preston *et al.*, 1994; Thomsen *et al.*, 1994) UL31 (Chang and Roizman, 1993), UL52 (Klinedinst and Challberg, 1994), UL54 (Phelan *et al.*, 1993); and for HHV-6 protein products of genes U11 (Neipel *et al.*, 1991; Pellett *et al.*, 1993) U16 (Geng *et al.*, 1992; Nicholas and Martin, 1994), U27 (Chang and Balachandran, 1991; Zou *et al.*, 1994), U39 (Ellinger *et al.*, 1993), U48 (Liu *et al.*, 1993a,b; Qian *et al.*, 1993), U57 (Littler *et al.*, 1990), U73 (Inoue *et al.*, 1994), U82 (Liu *et al.*, 1993a,b), U89/90 (Martin *et al.*, 1991b; Schiewe *et al.*, 1994), U95 (Thomson *et al.*, 1991), U100 (Pfeiffer *et al.*, 1993), and DR7 (Thompson *et al.*, 1994).


a
b
IE
IE-B
ME
IE-A
I
II
III
IV
V
VI
VII
LT1
DR1
RJ1
US22
US26
RS1
R1
R2
R3

They both induce neutralizing antibodies, the gB antibodies requiring complement (Foa-Tomasi *et al.,* 1991; Chou and Marousek, 1992; Ellinger *et al.,* 1993; Liu *et al.,* 1993a,b; Qian *et al.,* 1993), while gH antibodies can neutralize infectivity as well as inhibit cell to cell spread by cell fusion (Foa-Tomasi *et al.,* 1991; Liu *et al.,* 1993a). Glycoprotein U18 and its homologue HCMV UL37 appears to be betaherpesvirus specific. In HCMV this gene is spliced and encodes an immediate-early glycoprotein. (Kouzarides *et al.,* 1988; Nicholas and Martin, 1994). In HHV-6 there is no evidence for splicing and the single ORF has all the characteristics of a class I glycoprotein. Two of the glycoproteins, U20 and U85, show similarity to cellular counterparts, which are members of the Ig superfamily (Williams and Barclay, 1988), both sharing a characteristic cysteine domain found initially in the constant region of Ig (U20: FASTA 112, chain C, Ig Epsilon; U85: FASTA 167, OX-2). Members of this family have roles in cell-to-cell contact via protein–protein interactions (Williams and Barclay, 1988) and it is postulated that expression of these glycoproteins will function in infected cell adhesion to uninfected lymphocytes or monocytes. Finally, U21, U22, and U23 appear to be unique to HHV-6. There are glycoproteins encoded in a similar location in HCMV, but they share no detectable similarity; further, they share no similarity among themselves. However, the adjacent U20 and U21, as well as the adjacent U22 and U23, encode proteins with similar predicted molecular weights (49 and 25 kDa, for each pair, respectively), and all are from the complementary strands indicating possible gene duplication events followed by divergence.

There are other genes which appear to encode membrane proteins, six with single and three with multiple membrane spanning regions. ORFs U24, U46, U47, U50, U82, and U83 have single membrane spanning regions. U82 is the conserved glycoprotein gL, which complexes with gH to form a heterodimer membrane glycoprotein (Liu *et al.,* 1993a,b). U46 has an N-terminal region of length and composition to be a class II membrane protein with uncleaved signal peptide (Von Heijne and Gavel, 1988; Gompels *et al.,* 1992). Membrane proteins of similar character are encoded at a similar locus in all herpesviruses so far examined (Table 3) (Barker and Roizman, 1992; Barnett *et al.,* 1992; Telford *et al.,* 1992). The HSV positional homologue codes for a virion protein (Barker and Roizman, 1992). U47 is present at a major site of rearrangement between herpesviruses (Figs. 3 and 4) and is only present in the betaherpesviruses where corresponding genes encode a glycoprotein with little sequence similarity. Both may be secreted or anchored at the N-terminus and have multiple sites for both N- and O-linked glycosylation (Gompels *et al.,* 1992). There is evidence that this glycoprotein varies *in vivo* between strains and *in vitro* during passage of the virus in tissue culture (Gompels *et al.,* 1993, 1994). U83 is of particular interest since it has some features of an intercrine cytokine: it is small (10 kDa), contains a dicysteine "CC" motif (CC-28X-C-15X-C like the chemokines CC-23X-C-13X-C), and is basic overall, with the potential to be secreted like other members of this cytokine family (Oppenheim *et al., 1991).*

Three multiple membrane spanning proteins are predicted from the coding sequences of U12, U51, and U72. Both U12 and U51 have similarity to all G-protein-coupled receptors on the current database, and these have only been identified in beta and gammaherpesviruses. U12 is most related to GCRs identified in HCMV, UL33 (FASTA 433), as well as US27, US28 (Chee *et al.,* 1990), and the lymphocyte-specific GCR induced by EBV, EBI1 (FASTA 245) (Birkenbach *et al.,* 1993). In this regard, HHV-6 infection has recently been shown to induce the EBI1 (Hasegawa *et al.,* 1994). HCMV US28 has been shown to function as a chemokine receptor (Neote *et al.,* 1993) and U12 also shows close relation to a type I chemokine receptor (FASTA 258). Both U12 and HCMV UL33 appear to have a small N-terminal exon with consensus donor/acceptor sites (Fig. 2; Table 1) and members of the inflammatory mediator GCRs also have a similarly placed exon (Gerard and Gerard, 1994). Comparisons of the spliced U12 to the GCR family show closer similarity and include an additional conserved cysteine and N-linked glycosylation site. U51 is more closely related to HCMV UL78 and the HVS gene74 GCR (FASTA 146 and 139, respectively) (Nicholas *et al., 1992*), than the other HCMV GCRs. HCMV UL78 was not noted previously as a GCR, although we show here (summarized in Table 3) it's relationship to HHV-6 U51 and the previously described HVS GCR (Nicholas *et al.,* 1992). Interestingly, U51 is more closely related to the cellular GCRs, opioid receptors (FASTA 204), than to its viral counterparts, although both are related to chemokine receptors. In contrast, the "integral membrane protein" U72, gM, another multiple membrane spanning glycoprotein, is more closely related to homologues in all herpesviruses examined (Lawrence *et al.,* 1995) (Table 3), but has little similarity to cellular proteins. In HSV and EHV-1, the gM homologue is a major component of the virus particle (Baines and Roizman, 1993; Pilling *et al.,* 1994).

A fourth multiple membrane protein is implied from

FIG. 3. Conserved gene organization between HHV-6 and HCMV. Genes which encode homologous amino acid sequences are shaded. HCMV is shown with the I-S (inverted S region) orientation. Gene "blocks" I–VII containing genes conserved in all herpesviruses (see the legend to Fig. 4 and Table 3) are marked in relative positions in HHV-6 and HCMV and show regions of divergence between other herpesviruses are also regions of divergence within the betaherpesvirus group. Preliminary analyses of immediate early expressing genes are indicated for HHV-6 as IE-A and IE-B, which are in similar relative positions to HCMV major IE (MIE) and additional IE regions. Long repetitive sequences are indicated as blocks and unique regions as solid lines.

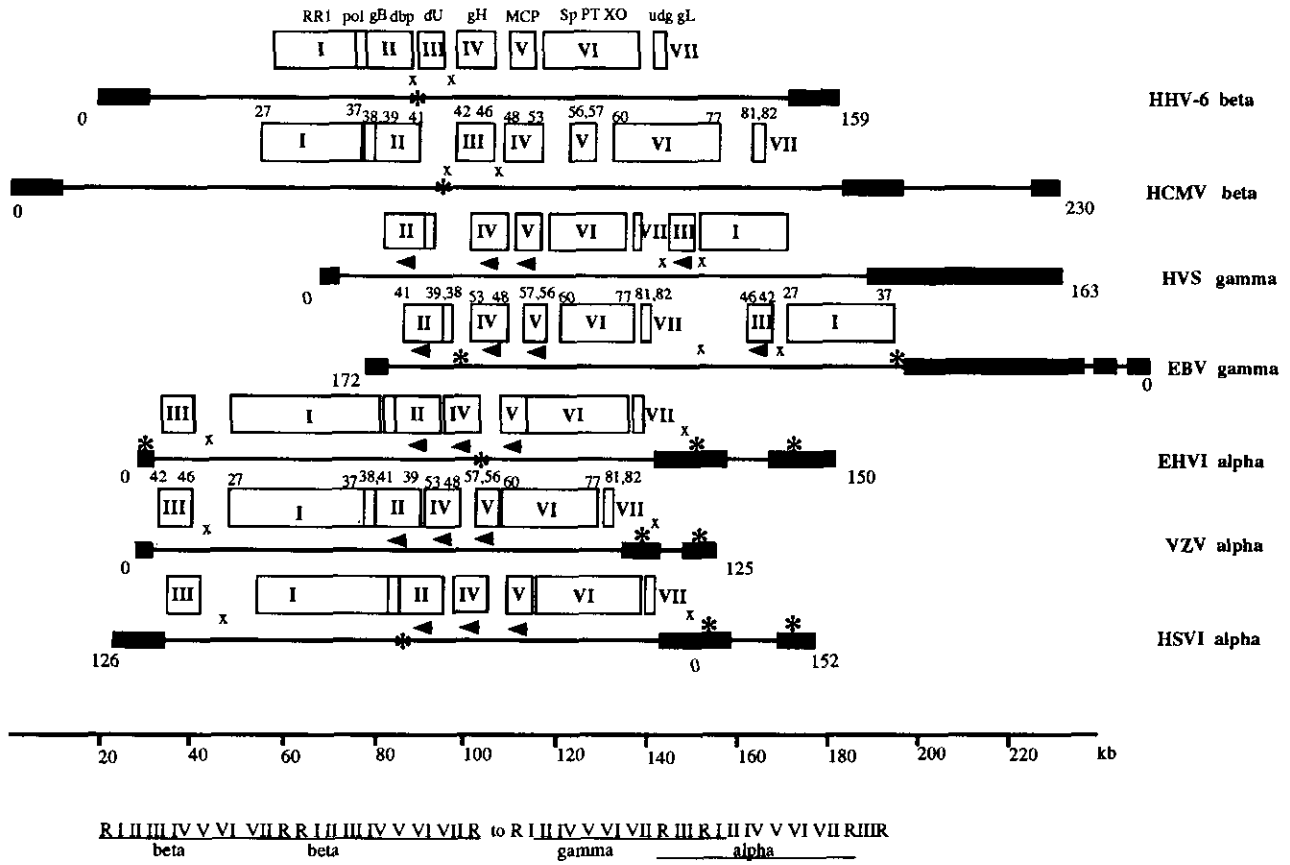


FIG. 4. Comparison of the gene organization in alpha, beta, and gammaherpesviruses as represented by the sequenced genomes of HSV-1, VZV, EHV-1 (alpha); EBV, HVS (gamma); and HCMV, HHV-6 (beta) (Baer *et al.*, 1984; Davison and Scott, 1986; McGeoch *et al.*, 1988; Chee *et al.*, 1990; Albrecht *et al.*, 1992; Telford *et al.*, 1992). Gene blocks I–VII are shown and represent regions with genes in similar orientation and encoding homologous amino acid sequences as shown in Tables 1 and 3 and also in Fig. 3. Inverted gene blocks are indicated by arrowheads. Abbreviations are as described in the legend Fig. 2. The "x" marks regions of divergence adjacent to the rearranged block III (see text). The star indicates origins of lytic replication. Long repeated regions in each genome are shown by solid boxes. Below the scale a diagram summarizes rearrangements between lineages. Genomes are shown adjacent as in a concatemer. Block III moves into the repeat regions followed by separate aberrant cleavage/packaging events to produce gamma or alpha lineages.

recent cDNA analyses. U100 encodes a single identifiable transmembrane region, but it is one of 10 exons encoding glycoprotein gp82/105 with other potential transmembrane regions (Pfeiffer *et al.*, 1993) (B. Chandran and B. J. Thomson, in press). Neutralizing monoclonal antibodies have been isolated which are specific for a peptide encoded by U100. This indicates that gp82/105 is another glycoprotein involved in virus infection; it also appears to be HHV-6 specific (Balachandran *et al.*, 1989; Pfeiffer *et al.*, 1993).

IE proteins and "transactivation"

Membrane proteins play a role in virus spread and are a first line in determining viral tropisms. IE proteins on the other hand control the "temporal cascade" of gene expression and are a second line in determining the cellular specificity of infection. The HHV-6 GCRs may play a role in cellular activation to induce viral gene expression, but it is not known if these are expressed early. U18 is a homologue of the HCMV UL37 IE glycoprotein shown to be expressed as an immediate early protein (Kouzarides *et al.*, 1988) and may play a similar role in cell activation. At least seven genes have been defined which may serve as transcriptional activators based on *in vitro* CAT assays: DR7, U16/U17, U18/U19, U27, U86/U87, U89, and U94 (Martin *et al.*, 1991b; Geng *et al.*, 1992; Nicholas and Martin, 1994; Thompson *et al.*, 1994; Zhou *et al.*, 1994; Thomson *et al.*, 1994b) (S. Holmes and J. Nicholas, unpublished results). One of these, spliced U89/U90, has been reported to be expressed as an IE gene and is transcribed in the absence of viral protein synthesis. U91 is similarly expressed, although it becomes more abundant at later stages postinfection (Schiewe *et al.*, 1994). There are preliminary data that indicate that IE transcripts are also made from the U18/U19 and U86/U87 loci (S. Holmes and J. Nicholas, unpublished results). These findings, together with the homologies of U16/U17, U18, U19, and U86 to HCMV genes within either the UL36-38 or the major IE loci (Chee *et al.*, 1990; Colberg-Poley *et al.*, 1992; Nicholas, 1994; Nicholas and Martin, 1994) suggest that the equivalent loci in HHV-6 are functionally analogous. We have therefore assigned the U86–U89 and U16–U19 loci the names IE region A (IE-A) and IE

TABLE 2

HHV-6 Shares with HCMV and MCMV the HCMV US22 Gene Family[a]

name	DR2	DR7	U2	U3	U7	U8	U16	U25	U95	'CC' motif
DR2	-		77	81	77		95	77	70	-
DR7		-		95						+
U2	77		-	144		77	91	203	111	+
U3	81	95	144	-		100	185			+
U7	77				-	97		96		+/-
U8			77	100	97	-		93	82	-
U16	95		91	185			-			+
U25	77		203		96	93		-	120	+
U95	70		111			82	89	120	-	+/-
MCMVIE2			108				163	120	284 P	+
HCMVUL23			P	137			95			+
HCMVUL24				137 P	92			201		+
HCMVUL28					189 P				106	-
HCMVUL29			113			131 P			90	+
HCMVUL36							P	162		+
HCMVUL43								94 P		+
HCMVIRS1									P	-
HCMVUS22		81 P	101				98			+
HCMVUS23				99			121			+
HCMVUS24				112		124	104			+
HCMVUS26	144 P						108			+
HCMVTRS1									P	-

[a] FASTA scores above 70 are shown. "CC" motif as described in text. P marks the positional homologue of each gene.

region B (IE-B), respectively (Fig. 2). Region A corresponds to the major IE1 and 2 regions of HCMV (Chee *et al.,* 1990); U86 is homologous to IE2 and U89 is unique to HHV-6 but is a positional homologue of IE1. Based on encoded amino acid sequence comparisons, other regions may be involved in transcriptional activation, including U41, U54, and U95. U41 is a member of the only conserved herpesvirus gene involved in "transactivation." The HCMV homologue, UL69, has been shown to transactivate gene expression *in vitro* (Winkler *et al.,* 1994), but studies on the HSV homologues also show a role in transcriptional termination, and snRNP distribution (McLauchlan *et al.,* 1992; Phelan *et al.,* 1993). U54 has similarity to a gene duplication in HCMV, UL82 and UL83, and UL82 has recently been shown as a transcriptional activator in the tegument (Liu and Stinski, 1992). U95 has sequence similarity and is a positional homologue of MCMV IE2 (Messerle *et al.,* 1991) which is the third gene in the major immediate early region of MCMV. In HCMV this region is more divergent and it may correspond to the transactivating genes, TRS and IRS in the adjacent S region (Fig. 3) (Stasiak and Mocarski, 1992). U95 is a member of the US22 family, as well as two functionally defined transcriptional activators, DR7 and U16. Thus, it is possible that other US22 family members may also be involved in control of gene expression (Table 2). A number of proteins have domains with SR repeats and other features of SR proteins, some of which can bind RNA, are involved in RNA processing and are translocated to the nucleus (Birney *et al.,* 1993; Gui *et al.,* 1994). The R1 repeats in U86 (Nicholas, 1994), the HCMV IE2 homologue, are a clear example, as well as the minor repeats in DR8 and U1 (Gompels and Macaulay, 1995).

DNA replication proteins

As in the host, transcription and replication factors may have overlapping roles as multifunctional proteins in a replication complex or to enhance expression of replication genes. U27 encodes a protein which functions in transactivation (Zhou *et al.,* 1994) but may also function in replication as it is the positional homologue of the DNA polymerase processivity factor (Tables 1 and 3; Fig. 3). Initial studies of HSV-1 replication, using a transient-transfection assay, demonstrated that seven genes were necessary and sufficient for replication (Challberg and Kelly, 1989). These are all conserved in HHV-6. In addition to U27, there is U38 (DNA Polymerase), U41 (single-stranded DNA binding protein), U43, U74, U77 (helicase/primase complex; U43 primase, U77 helicase), and U73 (origin binding protein). In HCMV a similar transient-transfection assay identified additional loci encoding *trans*-acting factors for replication, although HCMV lacks a detectable homologue of U73 (Pari and Anders, 1993). These included transcriptional activators, major IE1, IE2, UL36, and IRS1/TRS1, which have homologues in HHV-6 as described above. The polymerase homologue, U38,

TABLE 3

Homologues of HHV-6 Genes in Sequenced Herpesvirus Genomes Representing Alpha- (HSV-1, VZV, EHV-1), Beta-(HCMV), and Gamma-herpesviruses (EBV, HVS)[a,b]

Gene Block	HHV6	HHV6:HHV6	:HCMV	:HVS	:EBV	:EHV1	:VZV	:HSV1
	DR1	DR6 : 131						
	DR2	U16 : 95	HCMVUS26 : 144					
	DR6	DR1 : 131						
	DR7	U3 : 95	HCMVUS22 : 81					
	U2	U25 : 203	HCMVUL29 : 113					
	U3	U16 : 185	HCMVUL24 : 137					
	U4	U5 : 154	HCMVUL27 : P					
	U5	U4 : 151	HCMVUL27 : P					
	U7	U8 : 97	HCMVUL28 : 189					
	U8	U3 : 100	HCMVUL29 : 131					
	U10		HCMVUL31 : 397					
	U11		HCMVUL32 : 293					
	U12	U51 : 95	HCMVUL33 : 433	GENE74 : P	EBI1* : 245			
	U14		HCMVUL25 : 293			EHV63? : 111		IE110?
	U16	U3 : 185	HCMVUS23 : 121					
	U18		HCMVUL37 : 112					
	U19		HCMVUL38 : 214					
	U25	U2 : 203	HCMVUL24 : 201					
I	U27		HCMVUL44 : 626	GENE59 : P	BMRF1 : P	EHV18 : P	VZV16 : P	UL42 : P
I	U28		HCMVUL45 : 482	GENE61: 280	BORF2: 181	EHV21: 313	VZV19: 292	UL39: 271
I	U29		HCMVUL46 : 162	GENE62 : P	BORF1 : 97	EHV22 : P	VZV20 : P	UL38 : P
I	U30		HCMVUL47 : 1047	GENE63 : 208	BOLF1 : P	EHV23 : P	VZV21 : P	UL37? :
I	U31		HCMVUL48 : 449	GENE64 : 233	BPLF1 : 122	EHV24 : 145	VZV22 : 154	UL36 : 160
I	U33		HCMVUL49 : 669	GENE66 : 235	BFRF2 : 196	EHV25 : P	VZV23 : P	UL35 : P
I	U34		HCMVUL50 : 464	GENE67 : P	BFRF1 : 131	EHV26 : P	VZV24? : P	UL34? : P
I	U35		HCMVUL51 : 243			EHV27 : 110	VZV25 : 80	UL33 : 82
I	U36		HCMVUL52 : 578	GENE68 : 315	BFLF1 : 203	EHV28 : 195	VZV26 : 163	UL32 : 161
I	U37		HCMVUL53 : 670	GENE69 : 238	BFLF2 : 236	EHV29 : P	VZV27 : 106	UL31 : 147
II	U38		HCMVUL54 : 505	GENE9 : 1954	BALF5 : 1998	EHV30 : 909	VZV28 : 865	UL30 : 874
II	U39		HCMVUL55 : 1771	GENE8 : 1156	BALF4 : 759	EHV33 : 974	VZV31 : 635	UL27 : 663
II	U40		HCMVUL56 : 1107	GENE7 : 693	BALF3 : 597	EHV32 : 297	VZV30 : 395	UL28 : 368
II	U41		HCMVUL57 : 1392	GENE6 : 456	BALF2 : 857	EHV31 : 201	VZV29 : 429	UL29 : 233
III	U42		HCMVUL69 : 514	GENE57 : P	BMLF1 : P	EHV5 : P	VZV4 : P	UL54 : P
III	U43		HCMVUL70 : 588	GENE56 : 534	BSLF1 : 650	EHV7 : 114	VZV6 : 306	UL52 : 160
III	U44		HCMVUL71 : 304	GENE55 : P	BSRF1 : P	EHV8 : P	VZV7 : P	UL51? :
III	U45		HCMVUL72 : 383	GENE54 : P	BLLF2 : P	EHV9 : P	VZV8 : P	UL50 : P
III	U46		HCMVUL73 : 172	GENE53 : 143	BLRF1 : 157	EHV10 : P	VZV9A : P	UL49A : P
	U47		HCMVUL74 : P	GENE51? : 101				
IV	U48		HCMVUL75 :795	GENE22 : 298	BXLF2 : 267	EHV39 : 106	VZV37 : 92	UL22 : P
IV	U49		HCMVUL76 : 389	GENE20 : 223	BXRF1 : 209	EHV37 : 130	VZV35 : 161	UL24 : 106
IV	U50		HCMVUL77 : 845	GENE19 : 194	BVRF1 : 360	EHV36 : 350	VZV34 : 351	UL25 : 329
	U51	U12:95	HCMVUL78 : 146	GENE74 : 139				
IV	U52		HCMVUL79 : 513	GENE18 : 281	BVRF1.5a/b			
IV	U53		HCMVUL80 : 456	GENE17 : 365	BVRF2 : 380	EHV35 : 226	VZV33 : 98	UL26 : 244
	U54		HCMVUL82 : 90					
	U55		HCMVUL84 : 170					
V	U56		HCMVUL85 : 766	GENE26 : 222	BDLF1 : 234	EHV43 : 118	VZV41 : 99	UL18 : 103
V	U57		HCMVUL86 : 2143	GENE25 : 972	BcLF1 : 957	EHV42 : 540	VZV40 : 425	UL19 : 613
	U58		HCMVUL87 : 989	GENE24 : 505	BcRF1 : 514			
	U59		HCMVUL88 : 183					
VI	U60		HCMVUL89EX2 : 1072	GENE29b : 629	BDRF1 : 600	EHV44 : 664	VZV42 : 634	UL15EX2 : 694
	U62		HCMVUL91 : 105					
	U63		HCMVUL92 : 509	GENE31 : 146	BDLF4 : 199			
VI	U64		HCMVUL93 : 120	GENE32 : 133	BGLF1 : P	EHV45 : P	VZV43 : P	UL17 : P
VI	U65		HCMVUL94 : 538	GENE33 : 226	BGLF2 : 227	EHV46 : 121	VZV44 : P	UL16 : 114
VI	U66		HCMVUL89EX1 : 841	GENE29a : 359	BGRF1 : 359	EHV47 : 328	VZV45 : 297	UL15EX1 : 302
VI	U67		HCMVUL95 : 260	GENE34 : 264	BGLF3 : P	EHV48 : P	VZV46 : P	UL14 : P
	U68		HCMVUL96 : 127	GENE35 : P	BGLF3.5 : P			
VI	U69		HCMVUL97 : 530	GENE36 : 167	BGLF4 : 141	EHV49 : 131	VZV47 : 147	UL13 : 127
VI	U70		HCMVUL98 : 374	GENE37 : 406	BGLF5 : 420	EHV50 : 158	VZV48 : 108	UL12 : 173
VI	U72		HCMVUL100 : 887	GENE39 : 449	BBRF3 : 422	EHV52 : 259	VZV50 : 234	UL10 : 178
	U73					EHV53 : 336	VZV51 : 370	UL9 : 436
VI	U74		HCMVUL102 : 115	GENE41 : P	BBLF3 : P	EHV54 : P	VZV52 : P	UL8 : P
VI	U75		HCMVUL103 : 288	GENE42 : 134	BBRF2 : P	EHV55 : 116	VZV53 : 109	UL7 : 104
VI	U76		HCMVUL104 : 528	GENE43 : 698	BBRF1 : 521	EHV56 : 211	VZV54 : 257	UL6 : 166
VI	U77		HCMVUL105 : 816	GENE44 : 887	BBLF4 : 1206	EHV57 : 1021	VZV55 : 744	UL5 : 749
	U79		HCMVUL112 : 132					
	U80		HCMVUL113 : 95					
VII	U81		HCMVUL114 : 708	GENE46 : 595	BKRF3 : 555	EHV61 : 487	VZV59 : 506	UL2 : 441
VII	U82		HCMVUL115 : 244	GENE47 : P	BKRF2 : P	EHV62 : P	VZV60 : P	UL1 : P
	U84		HCMVUL117 : 170					
	U86		HCMVUL122 : 284					
	U95	U2 : 111	HCMVUL28 : 106					
	'total'/ no : avg		**31374 / 68 : 461**	**15209 / 39 : 390**	**13610 / 31 : 439**	**8671 / 27 : 321**	**8057 / 26 : 310**	**7887 / 25 : 315**
	'core'/no : avg		**17227 / 21 : 820**	**11787 / 21 : 561**	**11731 / 21 : 558**	**7771 / 21 : 370**	**7300 / 21 : 347**	**7115 / 21 : 339**
	total kb, %GC	159, 43%	230, 57%	164, 41%	172, 60%	150, 57%	125, 46%	152, 68%
	core kb, %GC	86, 40%	109, 59%	89, 33%	109, 58%	104, 55%	99, 44%	106, 67%

shows the closest relationship to the host polymerase delta (30% identity over 733 amino acids), like other herpesvirus polymerases, rather than polymerase alpha as originally thought (Yang *et al.*, 1992). Polymerase delta has recently been shown in an SV40 model system to replicate both leading and lagging strands, while polymerase alpha has a role in DNA priming (Waga and Stillman, 1994).

Two new replication specific loci have been identified for HCMV by the transient assay, UL84 and UL112-113 (Pari and Anders, 1993), which encode proteins with low similarity to HHV-6 U55 and U79-U80. Although these loci may indicate features of replication that are unique to betaherpesviruses, previous analyses show that they are highly divergent and functional assays are required to determine their role (Nicholas, 1994). HHV-6 may have some individual features to its replication strategy in that it has a homologue, U94, to the parvovirus rep protein, a multifunctional protein with a role in priming DNA replication (Thomson *et al.*, 1991). Initial studies suggest that U94 complements the parvovirus rep protein (Thomson *et al.*, 1994b).

Like other herpesviruses, HHV-6 encodes a number of proteins involved in DNA repair and nucleotide metabolism. Both HHV-6 and HCMV lack a thymidine kinase (ORF deletion between U48/U49 and corresponding HCMV UL75/76) although HCMV encodes another kinase which can modify nucleosides, UL97 (Littler *et al.*, 1992; Sullivan *et al.*, 1992) and the HHV-6 homologue is U69. This kinase is conserved in alpha, beta, and gamma herpesviruses (Chee *et al.*, 1989; Lawrence *et al.*, 1990), but in HCMV mutations in this gene and the polymerase confer resistance to ganciclovir (Littler *et al.*, 1992; Sullivan *et al.*, 1992, 1993). As HHV-6 has some similar drug sensitivities to HCMV, these genes may also be sites for drug resistance in HHV-6 (AkessonJohansson *et al.*, 1990). U28 and U45, and their HCMV homologues (Table 1), may be further betaherpesvirus specific features involved in replication as they have diverged from their respective protein families. U28, the large subunit of ribonucleotide reductase (RR1) homologue, has no adjacently encoded small subunit as in the alpha- and gamma-herpesviruses. Studies on HSV show that these subunits can bind to make the functional enzyme (Conner *et al.*, 1993). It is not known whether RR1 of betaherpesviruses binds to cellular components or has a different function. Similarly, U45 and HCMV UL72, the putative betaherpesvirus dUTPase homologues (Gompels *et al.*, 1992), are substantially divergent from the rest in this group, and may have a different specificity or function as more distant members of this protein family (McGeoch, 1990).

DNA packaging and virus assembly proteins

There are a number of proteins with a role in packaging of the nascent replicated DNA, some of which group as structural proteins. U53 encodes the conserved herpesvirus protease/assemblin which can cleave itself and the internal in-frame, U53a, scaffolding protein homologue. The scaffolding protein acts as a frame for DNA packaging and capsid formation, but is not packaged itself in the mature virion. Temperature sensitive mutants of the HSV protease, UL26, and in-frame scaffold, UL26.5, do not have proteolytic activity or package DNA at the nonpermissive, temperature indicating that protease activity is necessary for packaging and this is supported by studies on null mutants (Addison *et al.*, 1984; Desai *et al.*, 1994; Gao *et al.*, 1994). Studies on HHV-6 U53 indicate that the expressed protein can cleave itself and has two sites, release and maturation (C. Parry and U. A. Gompels, in preparation), which match consensus sequences determined for HCMV and SCMV by site-directed mutagenesis (Welch *et al.*, 1993; Jones *et al.*, 1994). U29, U30, homologues of HCMV UL46, UL47 (Chee *et al.*, 1990), and the homologue of the late spliced conserved gene, U60/66, have also been implicated in DNA packaging and capsid assembly (Poon and Roizman, 1993).

In addition to genes encoding virion glycoproteins such as gB and gH/gL (U39 and U48/U82), there are genes for herpesvirus-conserved tegument and capsid proteins that are also found in HHV-6. Tegument proteins include the major antigenic phosphoprotein, pp100, U11, homologue of HCMV pp150 (Neipel *et al.*, 1992; Pellett *et al.*, 1993), large tegument protein, U48, and possible HCMV pp65 homologue, U54. Capsid proteins include the major capsid protein, U57 (Littler *et al.*, 1990), as well as probable capsid proteins U33, U56 and virion proteins U34, U36, U50 (Chee *et al.*, 1990; Telford *et al.*, 1992). HSV-1 free capsid assembly has been observed using baculovirus recombinants of HSV UL18, 19, 35, 38, 26, and 26.5 (Newcomb *et al.*, 1994; Preston *et al.*, 1994; Thomsen *et al.*, 1994). HHV-6 encodes the homologues

TABLE 3—*Continued*

[a] HHV-6: HHV-6 are the results of self comparison. P marks positional/probable homologues. *marks a cellular gene induced by EBV, EBI1. ? marks indicate a tentative similarity based mainly on compositional biases. Gene blocks I–VII are genes conserved in alpha, beta, and gamma herpesviruses as shown (amino acid sequence similarity and/or positional homologues) which are rearranged relative to each other in each lineage, see Fig. 4.

[b] Sums below the columns show total FASTA scores, the number of matching gene products, and the average of these. Also shown is the "core," this is the gene products with detectable similarity in all these herpesviruses. The total kb of the respective genomes and G + C composition are shown, as well as that region corresponding to the "core" genes.

of all of these in U56, U57, U29, U33, U53, and U53a (Table 3) and these genes also correspond to those defined in HSV as encoding the components of the virus capsid (Davison *et al.*, 1992).

Betaherpesvirus relationship

The gene layouts of HHV-6 is very similar to that of HCMV but different from the alpha and gammaherpesviruses genomes which have been examined to date (Figs. 3 and 4). As described above, the HHV-6 proteins are also most closely related to HCMV proteins and this is easily shown by totalling the FASTA scores for the "core" conserved proteins and comparing to the scores of HCMV, EBV, EHV-1, VZV, and HSV. HHV-6 is most closely related not only in overall number of protein sequences matching, but also in the total score, as well as in the "core" comparison score (Table 3). Further, it is next related to the gammaherpesviruses and then least to all the alphaherpesviruses. This relationship can be summarized by taking the average FASTA score for "core" conserved genes (Table 3): HCMV 820, HVS 561, EBV 558, EHV-1 370, VZV 347, HSV-1 339. Most of the HHV-6 genes, 67%, have counterparts in HCMV detected either by sequence similarity or as positional homologues (Fig. 2; Tables 1, 2, and 3). HHV-6 specific genes with no detectable herpesvirus homologue include the following: the telomeric repeat ORFs LT1, LJ1, RJ1; the R2 ORF, U88; the glycoproteins U20, U21, U22, U23, U24 (although there are glycoproteins in this general region in HCMV they are unrelated); the OX-2 homologue, U85; the spliced glycoprotein gp82/105, U100; the parvovirus rep homologue, U94; the immediate early proteins U89/U90, U91; the DR ORFs DR1, DR3, DR4, DR5, DR8; as well as U1, U6, U9, U13, U15, U26, U32, U61, U78, U83, U92, U93, U96-99. Most of these are on either side of the genome and include regions with splicing; therefore, it is possible that some of these ORFs defined here as genes are in fact exons.

The betaherpesvirus specific genes, defined as those encoded by both HCMV and HHV-6, are also located mostly on either side of the genome. These include the gene families US22 (Table 2), DR1/DR6, U4/U5; structural and glycoprotein genes U11, U18, U47, U54; IE genes U16/17, U18, U19, U86; replication genes U55, U79, U80; and genes U10, U14, U59, U62. There are several genes which seem to be restricted to beta and gammaherpesviruses and may function as some common characteristic of these two lineages; these include U52, U58, U63, and U68. The GCR homologues may also be part of this group in that U51 is related to the HVS GCR gene 74 (Nicholas *et al.*, 1992) and although EBV does not encode one, it induces a GCR similar to U12 (Birkenbach *et al.*, 1993). Finally, there are at least two genes which show a relationship to alphaherpesviruses which are either lacking in gammaherpesviruses or divergent. These are the origin binding protein, U73, (also not found in HCMV) and a gene with unknown function, U35. Therefore, although the relationships between herpesviruses are clear from overall protein sequence comparisons (Table 3), there are some mosaic genes which may confer properties on individual viruses which are different from the entire group. Another example of such a gene is the thymidylate synthase in HVS and VZV (Honess *et al.*, 1986; Thompson *et al.*, 1987; Albrecht *et al.*, 1992).

Genetic rearrangements and genome evolution

An examination of the relative order of conserved genes contained in gene "blocks" in the sequenced herpesviruses suggest some of the major events contributing to the divergence in each lineage (Fig. 4). In terms of compactness and organization of the "core" conserved genes, it could be argued that as HHV-6 is the smallest (86 kb; Table 3), with few gaps between the gene blocks, that of the sequenced herpesviruses, it is more closely related to a "progenitor" herpesvirus which gave rise to the alpha and gammaherpesvirus lineages. Recent studies on dinucleotide relative abundance distances on available herpesvirus nucleotide sequences by Karlin and colleagues (1994b) have proposed an HHV-6 like virus as a "progenitor" herpesvirus, but further studies on more samples are required. The concordance between relationships determined by the nucleotide and protein sequence data is being further explored. Studies on fish herpesviruses show a completely different lineage (Davison, 1992) from alpha, beta, or gammaherpesviruses, and it may be that these three groups are only found in higher vertebrates such as birds or mammals. McGeoch and Cook (1994) have proposed from studies of alphaherpesvirus proteins that there is cospeciation with the host, the virus evolving at a rate of one or two magnitudes greater. This would be consistent with the divergence observed between fish and bird or mammalian herpesviruses (Davison, 1992; Karlin *et al.*, 1994b). It could be countered that there is no directionality to contraction and expansion; therefore, a larger "progenitor" such as EBV or HCMV (109 kb "core;" Table 3), rather than an HHV-6-like virus would be equally likely. However, these larger layouts are filled with repeated DNA or repeated gene families and it may be that these are later events or packaging requirements. The A+T rich "core" virus genomes are smaller than their G+C rich counterparts (Table 3), and it has been suggested that the G+C rich repeats may provide continual compositional biases (Honess, 1984). One example of these repeats would be the G+C rich pac1/2 signals for DNA packaging/cleavage located in terminal repetitive sequences (Honess, 1984; Deiss *et al.*, 1986; Gompels and Macaulay, 1995; Thomson *et al.*, 1994a).

The simplest model to account for the genetic rearrangements observed between alpha, beta, and gammaherpesvirus lineages as shown in Fig. 4, is that betaherpesviruses gave rise to the alpha and gamma-

herpesviruses and that A+T rich versions in each lineage are earlier forms. A minimum of three major events are required for these rearrangements over time. (i) The gene block III (flanked by ori–lyt) moves into the DR (possibly involving duplication). (ii) Aberrant DNA packaging or cleavage occurs, taking one genome-length from a concatemeric or circular form. Starting at block II and continuing to block I forms the gammaherpesvirus lineage, while starting at block III and continuing to block VII, forms the alphaherpesvirus lineages (summarized in Fig. 4, as shown for two adjacent genomes in a concatemer). (iii) Gene block inversions occur. These events are followed, accompanied, or preceded by repetitive sequence expansions or contractions and sequence drift.

There are at least five general observations which support the above model. (i) Regions of major lineage divergence are located where the block III region moves (marked by an X in Fig. 4), for example the EBNAs and gp350/220 in EBV, Vmw65 and gC in HSV. (ii) The origin of lytic replication seems involved in the transposition, in that where block III moves to or where it has moved from is marked by an ori. (iii) Regions containing ori–lyt can be duplicated intragenomically (Dewhurst *et al.*, 1993) as well as extragenomically in defective replicating subunits (Deiss *et al.*, 1986; Deiss and Frenkel, 1986). (iv) Transposition or recombination/deletion of ori–lyt to juxtaposition with *cis*-acting DNA packaging signals in the repeats has been observed for a number of families of defective replicating subunits originally described in HSV (Locker *et al.*, 1982; Mocarski *et al.*, 1985). (v) Regions which contain block III in gamma- or alpha-herpesviruses are adjacent to a relatively larger "gap" region.

Similar to that observed here for the herpesvirus family, lineage specific rearrangement of genes occur in the host genome (O'Brien *et al.*, 1988; Sankoff *et al.*, 1990; Green *et al.*, 1993) and repetitive DNA can mediate these events giving rise to a "modular" evolution of the genome (Charlesworth *et al.*, 1994). Regions of repetitive DNA mark rearrangements between alpha and gammaherpesviruses (Davison and Taylor, 1987; Gompels *et al.*, 1988) and this is extended by the observation here of repeats at the ori and terminal regions in betaherpesviruses and their relationship to rearrangements (Fig. 4).

In general, cellular gene homologues which are "inserted" in the genome are homogenized such that no detectable nucleotide sequence similarity remains (presumably as an ancient event involving an RNA intermediate as the genes do not contain cellular introns) and homology can only be detected by conserved amino acid sequence (Figs. 3 and 4; Table 3) (Chee *et al.*, 1990; Albrecht *et al.*, 1992). Likewise, relationships between herpesvirus lineages have been determined from encoded amino acid sequence similarity, the nucleotide sequence divergent beyond recognition. For example, HCMV and HHV-6 share little nucleotide sequence similarity, but have an average of 30% amino acid identity. In contrast, individual strains show considerable nucleotide sequence similarity in addition to their higher encoded amino acid sequence similarity.

Strain variation

Relationships between strains of HHV-6 have been described previously. These include close similarity in the center of the genome (to 1–5% variation) (Ablashi *et al.*, 1993; Aubin *et al.*, 1993; Gompels *et al.*, 1993) and increasing divergence at the ends (Neipel *et al.*, 1992; Pellett *et al.*, 1993; Chou and Marousek, 1994) adjacent to and in the DR repeats (Thomson *et al.*, 1994a). Examples of increased divergence between strains, including deletions and expansions, adjacent or in repetitive regions have been observed in HCMV (Chee *et al.*, 1990), EBV EBNA repeats and B958 deletion (Heller *et al.*, 1981; Parker *et al.*, 1990; Sample *et al.*, 1990), HSV "a" sequence (Locker and Frenkel, 1979), and EHV-1 repeats (Telford *et al.*, 1992). Comparisons of sequence presented here with 50 kb of sequence available from other strains support earlier preliminary findings. Overall strain variation in HHV-6 nucleotide sequences is more than that observed in the alphaherpesviruses HSV (3%) and EHV-1 (2%), but less than that observed in the beta and gammaherpesviruses, HCMV (5–11%) and EBV (4–36%) (Sakaoka *et al.*, 1987; Sample *et al.*, 1990; Chou and Dennison, 1991; Telford *et al.*, 1992; Gompels *et al.*, 1993; Fries *et al.*, 1994). Variation related to geographic origin has been observed for HSV and EBV (Sakaoka *et al.*, 1987; Packham *et al.*, 1993), and this has also been documented for human papillomaviruses, another persistent infection arising from close contact (Ong *et al.*, 1993). The predominant HHV-6 strain group appears to be variant B in the USA and this seems to apply to France, UK, Germany, and Japan (Dewhurst *et al.*, 1993; Drobyski *et al.*, 1993; Wilborn *et al.*, 1994), where the B variant averages 5% differences from the strains of the minor group, variant A, in studies of tissue samples, blood, or clinical isolates (Aubin *et al.*, 1993; Gompels *et al.*, 1993). As many A variants have been isolated from patients in or from Africa, it has been suggested that this may be the predominant strain group in that continent. Close association of strains with different populations could be possible given the efficient transmission, presumably maternal (via saliva or during birth), of this virus to most infants within their first year, and it's persistence in the host and host population (Cone *et al.*, 1993; Hall *et al.*, 1994). A complication of this effect may occur by blood-borne transmission during transfusion or reused needles in drug-abuse and by close contact or sexual transmission, as there is shedding from both salivary gland and cervical epithelia (as shown for other herpesviruses) (Fox *et al.*,

1990; Cone *et al.*, 1993; Chen *et al.*, 1994; Leach *et al.*, 1994).

Interactions with the immune system and HIV

HHV-6 can actively replicate and kill or abortively infect and remain latent in cells of the immune system which normally act to control infection. This must require an intricate balance for a virus which persists for the lifetime of the host. Examination here of the genes encoded by HHV-6 suggests some mechanisms. The virus encodes a number of proteins which have similarity to those with functions in the immune response and may be viewed as examples of molecular mimicry or immune evasion for both persistence and replication. These include members of the Ig superfamily and GCRs similar to chemokine receptors. The chemokines, a family of proinflammatory cytokines, act in migration and activation of leukocytes that mediate the inflammatory response (Oppenheim *et al.*, 1991; Gerard and Gerard, 1994). Interestingly the virus also encodes a product which may act as a CC chemokine to self stimulate these receptors and it could be speculated that the entire US22 family, which have a degenerate CC motif, may be processed to interact with the virus-encoded and -induced GCRs. The Ig-like glycoproteins may act after GCR-mediated migration/activation using their cysteine domains in cell attachment and/or in immune evasion, binding to similar receptors on lymphocytes/monocytes inducing anergy rather than activation. If HHV-6 is adapted to activating or repressing lymphocytes/monocytes, what may be the effects on cells infected with HIV?

During immunodeficiencies, including AIDS, HHV-6 can reactivate, and it is possible that HHV-6 in turn activates HIV. Cocultivation of these viruses *in vitro* can result in both enhancement or inhibition of HIV replication, depending on the conditions and cell types used (Lusso *et al.*, 1989; Carrigan *et al.*, 1990; Levy *et al.*, 1990). Indirect effects may be mediated by GCR signalling. Direct effects include transactivation of the HIV LTR promoter with a number of HHV-6 gene products: DR7, U16, U27, U89, and U94 (Martin *et al.*, 1991b; Geng *et al.*, 1992; Nicholas and Martin, 1994; Thompson *et al.*, 1994; Zhou *et al.*, 1994). Similar effects may be predicted for the HHV-6-related recent-isolate, HHV-7. Both viruses infect and kill helper T-lymphocytes and can cause exanthem subitum in primary infection (Frenkel *et al.*, 1990; Berneman *et al.*, 1992; Tanaka *et al.*, 1994). They have a similar gene organization, sharing protein sequence similarity (average 60%) and regions of limited nucleotide sequence identity (Berneman *et al.*, 1992) (J. Nicholas, unpublished results). HHV-7, like HIV, has been shown to utilize the CD4 antigen, as a cellular receptor (Lusso *et al.*, 1994) and infections with HHV-6 and HHV-7 can cause immune impairment by down regulation of surface receptors CD3 and CD4, respectively (Furukawa *et al.*, 1994). Thus, a close relationship between these T-cell tropic herpesviruses and HIV may exist and remains to be fully examined. The complete description of the gene complement of HHV-6 contributes to a further understanding of this virus and provides a blueprint to study its biology and pathogenic effects.

ACKNOWLEDGEMENTS

In global and local memory of the late Dr. R. W. Honess in whose Mill Hill, London NIMR laboratory this work was initiated and inspired. We also thank the support of the Wellcome Trust and the Medical Research Council (UK), and are grateful to Dr. B. Barrell and M-A. Rajastafan, Sanger Genome Centre, Hinxton, Cambridge, UK, for access to a prepublication version of the UNIX DIANA in Xanadu.

REFERENCES

Ablashi, D., Agut, H., Berneman, Z., Campadelli-Fiume, G., Carrigan, D., Ceccerini-Nelli, L., Chandran, B., Chou, S., Collandre, H., Cone, R., Dambaugh, T., Dewhurst, S., Diluca, D., Foa-Tomasi, L., Fleckenstein, B., Frenkel, N., Gallo, R., Gompels, U., Hall, C., Jones, M., Lawrence, G., Martin, M., Montagnier, L., Neipel, F., Nicholas, J., Pellett, P., Razzaque, A., Torrelli, G., Thomson, B., Salahuddin, S., Wyatt, L., and Yamanishi, K. (1993). Human herpesvirus-6 strain groups — A nomenclature. *Arch. Virol.* **129**, 1–4.

Addison, C., Rixon, F. J., Palfreyman, J. W., Ottara, M., and Preston, V. G. (1984). Characterisation of a herpes simplex virus type 1 mutant which has a temperature sensitive defect in penetration in cells and assembly of capsids. *Virology* **138**, 248–259.

AkessonJohansson, A., Harmenberg, J., Wahren, B., and Linde, A. (1990). Inhibition of human herpesvirus 6 replication by 9-(4-hydroxy-2-(hydroxymethyl)butyl)guanine (2HM-HBG) and other antiviral compounds. *Antimicrob. Agents Chemother.* **34**, 2417–2419.

Albrecht, J.-C., Nicholas, J., Biller, D., Cameron, K. R., Biesinger, B., Newman, C., Wittmann, S., Craxton, M. A., Coleman, H., Fleckenstein, B., and Honess, R. W. (1992). Primary structure of the Herpesvirus saimiri genome. *J. Virol.* **66**, 5047–5058.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. F. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Anders, D. G., Kacica, M. A., Pari, G., and Punturieri, S. M. (1992). Boundaries and structure of human cytomegalovirus oriLyt, a complex origin for lytic-phase DNA replication. *J. Virol.* **66**, 3373–3384.

Asano, Y., Yoshikawa, T., Kajita, Y., Ogura, R., Suga, S., Yazaki, T., Nakashima, T., Yamada, A., and Kurata, T. (1992). Fatal encephalitis/encephalopathy in primary human herpesvirus-6 infection. *Arch. Dis. Child.* **67**, 1484–1485.

Asano, Y., Yoshikawa, T., Suga, S., Yazaki, T., Kondo, K., and Yamanishi, K. (1990). Fatal fulminant hepatitis in an infant with human herpesvirus-6 infection. *Lancet* **335**, 862–863.

Aubin, J. T., Agut, H., Collandre, H., Yamanishi, K., Chandran, B., Montagnier, L., and Huraux, J. M. (1993). Antigenic and genetic differentiation of the two putative types of human herpesvirus 6. *J. Virol. Methods* **41**, 223–234.

Aubin, J. T., Collandre, H., Candotti, D., Ingrand, D., Rouzioux, C., Burgard, M., Richard, S., Huraux, J. M., and Agut, H. (1991). Several groups among human herpesvirus 6 strains can be distinguished by Southern blotting and polymerase chain reaction. *J. Clin. Microbiol.* **29**, 367–372.

Baer, R., Bankier, A. T., Biggin, M. D., Deininger, P. L., Farrell, P. J., Gibson, T. J., Hatfull, G., Hudson, G. S., Satchwell, S. C., Seguin, C., Tuffnell, P. S., and Barrell, B. G. (1984). DNA sequence and expression of the B95-8 Epstein-Barr virus genome. *Nature (London)* **310**, 207–211.

Baines, J. D., and Roizman, B. (1993). The UL10 gene of herpes simplex virus 1 encodes a novel viral glycoprotein, gM, which is present in the virion and in the plasma membrane of infected cells. *J. Virol.* **67**, 1441–1452.

Balachandran, N., Amelse, R. E., Zhou, W. W., and Chang, C. K. (1989). Identification of proteins specific for human herpesvirus 6-infected human T cells. *J. Virol.* **63,** 2835–2840.

Bankier, A. T., Weston, K. M., and Barrell, B. G. (1988). Random cloning and sequencing by the m13/dideoxynuculotide chain termination method. *Methods Enzymol.* **155,** 51–93.

Barker, D. E., and Roizman, B. (1992). The unique sequence of the herpes simplex virus 1 L component contains an additional translated open reading frame designated U(L) 49.5. *J. Virol.* **66,** 562–566.

Barnett, B. C., Dolan, A., Telford, E., Davison, A. J., and McGeoch, D. J. (1992). A novel herpes simplex virus gene (UL49A) encodes a putative membrane protein with counterparts in other herpesviruses. *J. Gen. Virol.* **73,** 2167–2171.

Berneman, Z. N., Ablashi, D. V., Li, G., Eger-Fletcher, M., Reitz, M., Jr., Hung, C. L., Brus, I., Komaroff, A. L., and Gallo, R. C. (1992). Human herpesvirus 7 is a T-lymphotropic virus and is related to, but significantly different from, human herpesvirus 6 and human cytomegalovirus. *Proc. Natl. Acad. Sci. USA* **89,** 10552–10556.

Biggin, M. D., Gibson, T. J., and Hong, C. F. (1983). Buffer gradient gels and S35 label as an aid to rapid DNA sequence determination. *Proc. Natl. Acad. Sci. USA* **80,** 3963–3965.

Bird, A. P. (1980). DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* **8,** 1499–1504.

Birkenbach, M., Josefsen, K., Yalamanchili, R., Lenoir, G., and Kieff, E. (1993). Epstein-Barr virus-induced genes: first lymphocyte-specific G protein-coupled peptide receptors. *J. Virol.* **67,** 2209–2220.

Birnboim, H. C., and Doly, J. (1979). A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7,** 1513–1523.

Birney, E., Kumar, S., and Krainer, R. (1993). Analysis of the RNA-recognition motif and RS an RGG domains: Conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Res.* **21,** 5803–5816.

Carrigan, D. R., Drobyski, W. R., Russler, S. K., Tapper, M. A., Knox, K. K., and Ash, R. C. (1991). Interstitial pneumonitis associated with human herpesvirus-6 infection after marrow transplantation. *Lancet* **338,** 147–149.

Carrigan, D. R., Knox, K. K., and Tapper, M. A. (1990). Suppression of human immunodeficiency virus type 1 replication by human herpesvirus-6. *J. Infect. Dis.* **162,** 844–851.

Challberg, M. D., and Kelly, T. J. (1989). Animal virus DNA replication. *Annu. Rev. Biochem.* **58,** 671–717.

Chang, C. K., and Balachandran, N. (1991). Identification, characterization, and sequence analysis of a cDNA encoding a phosphoprotein of human herpesvirus 6. *J. Virol.* **65,** 2884–2894.

Chang, Y. E., and Roizman, B. (1993). The product of the UL31 gene of herpes simplex virus 1 is a nuclear phosphoprotein which partitions with the nuclear matrix. *J. Virol.* **67,** 6348–6356.

Charlesworth, B., Sniegowski, P., and Stephan, W. (1994). The evolutionary dynamics of repetitive DNA in eukaryotes. *Nature (London)* **371,** 215–220.

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerney, R., Horsnell, T., Hutchinson, C. A., Kouzarides, T., Martignetti, J. A., Preddie, E., Satchwell, S. C., Tomlinson, P., Weston, K. M., and Barrell, B. G. (1990). Analysis of the protein coding content of the sequence of human cytomegalovirus strain AD169. *Curr. Top. Microbiol. Immunol.* **154,** 125–169.

Chee, M. S., Lawrence, G. L., and Barrell, B. G. (1989). Alpha-, beta- and gammaherpesviruses encode a putative phosphotransferase. *J. Gen. Virol.* **70,** 1151–1160.

Chen, M., Popescu, N., Woodworth, C., Berneman, Z., Corbellino, M., Lusso, P., Ablashi, D. V., and DiPaolo, J. A. (1994). Human herpesvirus 6 infects cervical epithelial cells and transactivates human papillomavirus gene expression. *J. Virol.* **68,** 1173–1178.

Chou, S., and Dennison, K. M. (1991). Analysis of interstrain variation in cytomegalovirus glycoprotein B sequences encoding neutralization-related epitopes. *J. Infect. Dis.* **163,** 1229–1234.

Chou, S., and Marousek, G. I. (1992). Homology of the envelope glycoprotein B of human herpesvirus-6 and cytomegalovirus. *Virology* **191,** 523–528.

Chou, S., and Marousek, G. I. (1994). Analysis of interstrain variation in a putative immediate-early region of human herpesvirus 6 DNA and definition of variant-specific sequences. *Virology* **198,** 370–376.

Colberg-Poley, A. M., Santomenna, L. D., Harlow, P. P., Benfield, P. A., and Tenney, D. J. (1992). Human cytomegalovirus US3 and UL36-38 immediate-early proteins regulate gene expression. *J. Virol.* **66,** 95–105.

Collins, J. F., and Coulson, A. F. W. (1987). *In* "Nucleic Acid and Protein Sequence Analysis: A Practical Approach." IRL Press, Oxford.

Cone, R. W., Hackman, R. C., Huang, M., Bowden, R. A., Meyers, J. D., Metcalf, M., Zeh, J., Ashley, R., and Corey, L. (1993). Human herpesvirus 6 in lung tissue from patients with pneumonitis after bone marrow transplantation. *New Engl. J. Med.* **329,** 156–161.

Cone, R. W., Huang, M., Ashley, R., and Corey, L. (1993). Human herpesvirus 6 DNA in peripheral blood cells and saliva from immunocompetent individuals. *J. Clin. Microbiol.* **31,** 1262–1267.

Conner, J., Furlong, J., Murray, J., Meighan, M., Cross, A., Marsden, H., and Clements, J. B. (1993). Herpes simplex virus type 1 ribonucleotide reductase large subunit: Regions of the protein essential for subunit interaction and dimerization. *Biochemistry* **32,** 13673–13680.

Corbellino, M., Lusso, P., Gallo, R. C., Parravicini, C., Galli, M., and Moroni, M. (1993). Disseminated human herpesvirus 6 infection in AIDS (18). *Lancet* **342,** 1242.

Costa, R. H., Draper, K. G., Kelly, T. J., and Wagner, E. K. (1984). An unusual spliced herpes simplex virus type 1 transcript with sequence homology to Epstein-Barr virus DNA. *J. Virol.* **54,** 317–328.

Davison, A. J. (1984). Structure of the genome termini of varicella-zoster virus. *J. Gen. Virol.* **65,** 1969–1977.

Davison, A. J. (1992). Channel catfish virus: A new type of herpesvirus. *Virology* **186,** 9–14.

Davison, A. J., and McGeoch, D. J. (1986). Evolutionary comparisons of the S segments in the genomes of herpes simplex virus type 1 and varicella-zoster virus. *J. Gen. Virol.* **67,** 597–611.

Davison, A. J., and Scott, J. E. (1986). The complete DNA sequence of varicella-zoster virus. *J. Gen. Virol.* **67,** 1759–1816.

Davison, A. J., and Taylor, P. (1987). Genetic relations between varicella-zoster virus and Epstein–Barr virus. *J. Gen. Virol.* **68,** 1067–1079.

Davison, M. D., Rixon, F. J., and Davison, A. J. (1992). Identification of genes encoding two capsid proteins (VP24 and VP26) of herpes simplex virus type 1. *J. Gen. Virol.* **73,** 2709–2713.

Deiss, L. P., Chou, J., and Frenkel, N. (1986). Functional domains within the a sequence involved in the cleavage-packaging of herpes simplex virus DNA. *J. Virol.* **59,** 605–618.

Deiss, L. P., and Frenkel, N. (1986). Herpes simplex virus amplicon: Cleavage of concatemeric DNA is linked to packaging and involves amplification of the terminally reiterated a sequence. *J. Virol.* **57,** 933–941.

Desai, P., Watkins, S. C., and Person, S. (1994). The size and symmetry of B capsids of herpes simplex virus type 1 are determined by the gene products of the UL26 open reading frame. *J. Virol.* **68,** 5365–5374.

Dewhurst, S., Dollard, S. C., Pellett, P. E., and Dambaugh, T. R. (1993). Identification of a lytic-phase origin of DNA replication in human herpesvirus 6B strain Z29. *J. Virol.* **67,** 7680–7683.

Dewhurst, S., McIntyre, K., Schnabel, K., and Hall, C. B. (1993). Human herpesvirus 6 (HHV-6) variant B accounts for the majority of symptomatic primary HHV-6 infections in a population of U.S. infants. *J. Clin. Microbiol.* **31,** 416–418.

Dolan, A., Arbuckle, M., and McGeoch, D. J. (1991). Sequence analysis of the splice junction in the transcript of herpes simplex virus type 1 gene UL15. *Virus Res.* **20,** 97–104.

Downing, R. G., Sewankambo, N., Serwadda, D., Honess, R. W., Crawford, D., Jarrett, R., and Griffen, B. E. (1987). Isolation of human lymphotropic herpesviruses from Uganda. *Lancet* **2,** 390.

Drobyski, W. R., Dunne, W. M., Burd, E. M., Knox, K. K., Ash, R. C., Horowitz, M. M., Flemenbery, N., and Carrigan, D. R. (1993). Human

herpesvirus-6 (HHV-6) infection in allogeneic bone marrow transplant recipients. I. Evidence of a marrow suppressive role for HHV-6 *in vivo*. *J. Infect. Dis.* **167,** 735–739.

Drobyski, W. R., Eberle, M., Majewski, D., and Baxter-Lowe, L. A. (1993). Prevalence of human herpesvirus 6 variant A and B infections in bone marrow transplant recipients as determined by polymerase chain reaction and sequence-specific oligonucleotide probe hybridization. *J. Clin. Microbiol.* **31,** 1515–1520.

Drobyski, W. R., Knox, K. K., Majewski, D., and Carrigan, D. R. (1994). Brief report: Fatal encephalitis due to variant B human herpesvirus-6 infection in a bone marrow-transplant recipient. *New Engl. J. Med.* **330,** 1356–1360.

Efstathiou, S., Lawrence, G. L., Brown, C. M., and Barrell, B. G. (1992). Identification of homologues to the human cytomegalovirus US22 gene family in human herpesvirus 6. *J. Gen. Virol.* **73,** 1661–1671.

Efstathiou, S., Minson, A. C., Field, H. J., Anderson, J. R., and Wildy, P. (1986). Detection of herpes simplex virus-specific DNA sequences in latently infected mice and in humans. *J. Virol.* **57,** 446–455.

Ellinger, K., Neipel, F., Foa-Tomasi, L., Campadelli-Fiume, G., and Fleckenstein, B. (1993). The glycoprotein B homologue of human herpesvirus 6. *J. Gen. Virol.* **74,** 495–500.

Ensoli, B., Lusso, P., Schachter, F., Josephs, S. F., Rappaport, J., Negro, F., Gallo, R. C., and Wong-Staal, F. (1989). Human herpesvirus-6 increases HIV-1 expression in co-infected T cells via nuclear factors binding to the HIV-1 enhancer. *EMBO J.* **8,** 3019–3027.

Fairfax, M. R., Schacker, T., Cone, R. W., Collier, A. C., and Corey, L. (1994). Human herpesvirus 6 DNA in blood cells of human immunodeficiency virus-infected men: Correlation of high levels with high CD4 cell counts. *J. Infect. Dis.* **169,** 1342–1345.

Foa-Tomasi, L., Boscaro, A., Di, G. S., and Campadelli-Fiume, G. (1991). Monoclonal antibodies to gp100 inhibit penetration of human herpesvirus 6 and polykaryocyte formation in susceptible cells. *J. Virol.* **65,** 4124–4129.

Fox, J. D., Briggs, M., Ward, P. A., and Tedder, R. S. (1990). Human herpesvirus 6 in salivary glands. *Lancet* **336,** 590–593.

Frenkel, N., Schirmer, E. C., Wyatt, L. S., Katsafanas, G., Roffmann, E., Danovich, R. M., and June, C. H. (1990). Isolation of a new herpesvirus from human CD4+ T cells. *Proc. Natl. Acad. Sci. USA* **87,** 748–752.

Fries, B. C., Chou, S., Boeckh, M., and Torok-Storb, B. (1994). Frequency distribution of cytomegalovirus envelope glycoprotein genotypes in bone marrow transplant recipients. *J. Infect. Dis.* **169,** 769–774.

Furukawa, M., Yasukawa, M., Yakushijin, Y., and Fujita, S. (1994). Distinct effects of human herpesvirus 6 and human herpesvirus 7 on surface molecule expression and function of CD4+ T cells. *J. Immunol.* **152,** 5768–5775.

Gao, M., MatusickKumar, L., Hurlburt, W., DiTusa, S. F., Newcomb, W. W., Brown, J. C., McCann, P., III, Deckman, I., and Colonno, R. J. (1994). The protease of herpes simplex virus type 1 is essential for functional capsid formation and viral growth. *J. Virol.* **68,** 3702–3712.

GeneticsComputerGroup (1991). Program Manual for the GCG Package Version 7. Madison, WI.

Geng, Y., Chandran, B., Josephs, S. F., and Wood, C. (1992). Identification and characterization of a human herpesvirus 6 gene segment that trans activates the human immunodeficiency virus type 1 promoter. *J. Virol.* **66,** 1564–1570.

Gerard, C., and Gerard, N. P. (1994). C5A Anaphylatoxin and its seven transmembrane segment receptor. *Annu. Rev. Immunol.* **12,** 775–808.

Gleeson, T. J., and Staden, R. (1991). An x-window and UNIX implementation of our sequence analysis package. *Comp. Applied Biosci.* **7,** 398.

Gompels, U. A., Carrigan, D. R., Carss, A. L., and Arno, J. (1993). Two groups of human herpesvirus 6 identified by sequence analyses of laboratory strains and variants from Hodgkin's lymphoma and bone marrow transplant patients. *J. Gen. Virol.* **74,** 613–622.

Gompels, U. A., Carss, A. L., Sun, N., and Arrand, J. R. (1992). Infectivity determinants encoded in a conserved gene block of human herpesvirus-6. *J. DNA Sequencing Mapp.* **3,** 25–39.

Gompels, U. A., Craxton, M. A., and Honess, R. W. (1988). Conservation of gene organization in the lymphotropic herpesviruses Herpesvirus saimiri and Epstein-Barr virus. *J. Virol.* **62,** 757–767.

Gompels, U. A., Luxton, J., Kehl, K. K., and Carrigan, D. R. (1994). Chronic bone marrow suppression in immunocompetent adult by human herpesvirus 6. *Lancet* **343,** 735–736.

Gompels, U. A., and Macaulay, H. A. (1995). Characterisation of human telomeric repeat sequences from human herpesvirus-6 and relationship to replication. *J. Gen. Virol.* **76,** 451–458.

Gompels, U. A., Nicholas, J., Martin, M. E. D., Lawrence, G., Jones, M., Thomson, B., Efstathiou, S., Teo, I., Craxton, M. A., Barrell, B., and Honess, R. W. (1992). The DNA sequence of human herpesvirus-6 strain U1102; a comparative analysis of gene organisation and coding content. *Int. Herp. Workshop Abs.* **17,** 43.

Green, P., Lipman, D. J., Hillier, L., States, D., Waterston, R., and Claverie, J. M. (1993). Ancient conserved regions in new gene sequences and the protein databases. *Science* **259,**(5102), 1711–1716.

Greider, C. W. (1993). Telomeres and telomerase in small eukaryotes. *Symp. Soc. Gen. Microbiol. (UK)* **50,** 31–42.

Gui, J.-F., Lane, W. S., and Fu, X.-D. (1994). A serine kinase regulates intracellular localization of splicing factors in the cell cycle. *Nature (London)* **369,** 678–682.

Hall, C. B., Long, C. E., Schnabel, K. C., Caserta, M. T., McIntyre, K. M., Costanzo, M. A., Knott, A., Dewhurst, S., Insel, R. A., and Epstein, L. G. (1994). Human herpesvirus-6 infection in children — A prospective study of complications and reactivation. *New Engl. J. Med.* **331,** 432–438.

Hamzeh, F. M., Lietman, P. S., Gibson, W., and Hayward, G. S. (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.

Harrington, L. A., and Greider, C. W. (1991). Telomerase primer specificity and chromosome healing. *Nature (London)* **353,** 451–454.

Hasegawa, H., Utsunomiya, Y., Yasukawa, M., Yanagisawa, K., and Fujita, S. (1994). Induction of G protein-coupled peptide receptor EBI 1 by human herpesvirus 6 and 7 infection in CD4+ T cells. *J. Virol.* **68,** 5326–5329.

Heller, M., Dambaugh, T., and Kieff, E. (1981). Epstein–Barr virus DNA. IX. Variation among viral DNAs from producer and nonproducer infected cells. *J. Virol.* **38,** 632–648.

Honess, R. W. (1984). Herpes simplex and 'the herpes complex': diverse observations and a unifying hypothesis. *J. Gen. Virol.* **65,** 2077–2107.

Honess, R. W., Bodemer, W., Cameron, K. R., Niller, H.-H., Fleckenstein, B., and Randall, R. E. (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.

Honess, R. W., Gompels, U. A., Barrell, B., Craxton, M. A., Staden, R., Chang, Y.-C., and Hayward, G. (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.

Inoue, N., Dambaugh, T. R., Rapp, J. C., and Pellett, P. E. (1994). Alphaherpesvirus origin-binding protein homolog encoded by human herpesvirus 6B, a betaherpesvirus, binds to nucleotide sequences that are similar to ori regions of alphaherpesviruses. *J. Virol.* **68,** 4126–4136.

Jones, M. D., and Teo, I. A. (1992). DNA sequence analysis of the human herpesvirus-6 genes encoding glycoprotein B and the major DNA binding protein. *Int. Herp. Workshop Abs* **17,** 62.

Jones, M. D., and Teo, I. A. (1993). Identification and analysis of the transport/capsid assembly protein (tp/cap) gene of human herpesvirus-6 (HHV6). *Virology* **197,** 449–454.

Jones, T. R., Sun, L., Bebernitz, G. A., Muzithras, V. P., Kim, H. J., Johnston, S. H., and Baum, E. Z. (1994). Proteolytic activity of human cytomegalovirus UL80 protease cleavage site mutants. *J. Virol.* **68,** 3742–3752.

Karlin, S., Doerfler, W., and Cardon, L. R. (1994a). Why is CpG suppressed in the genomes of virtually all small eukaryotic viruses but not in those of large eukaryotic viruses? *J. Virol.* **68,** 2889–2897.

Karlin, S., Mocarski, E. S., and Schachtel, G. A. (1994b). Molecular evolution of herpesviruses: Genomic and protein sequence comparisons. *J. Virol.* **68,** 1886–1902.

Kaye, J. F., Gompels, U. A., and Minson, A. C. (1992). Glycoprotein H of human cytomegalovirus (HCMV) forms a stable complex with the HCMV UL115 gene product. *J. Gen. Virol.* **73,** 2693–2698.

Kishi, M., Harada, H., Takahashi, M., Tanaka, A., Hayashi, M., Nonoyama, M., Josephs, S. F., Buchbinder, A., Schachter, F., Ablashi, D. V., Wong-Staal, F., Salahuddin, S. Z., and Gallo, R. C. (1988). A repeat sequence, GGGTTA, is shared by DNA of human herpesvirus 6 and Marek's disease virus. *J. Virol.* **62,** 4824–4827.

Klinedinst, D. K., and Challberg, M. D. (1994). Helicase-primase complex of herpes simplex virus type 1: A mutation in the UL52 subunit abolishes primase activity. *J. Virol.* **68,** 3693–3701.

Knox, K. K., and Carrigan, D. R. (1994). Disseminated active HHV-6 infections in patients with AIDS. *Lancet* **343,** 577–578.

Kondo, K., Kondo, T., Okuno, T., Takahashi, M., and Yamanishi, K. (1991). Latent human herpesvirus 6 infection of human monocytes/macrophages. *J. Gen. Virol.* **72,** 1401–1408.

Koszinowski, U. H., Keil, G. M., Volkmer, H., Fibi, M. R., Ebeling-Keil, A., and Munch, K. (1986). The 89,000-Mr murine cytomegalovirus immediate-early protein activates gene transcription. *J. Virol.* **58,** 59–66.

Kouzarides, T., Bankier, A. T., Satchwell, S. C., Preddie, E., and Barrell, B. G. (1988). An immediate early gene of human cytomegalovirus encodes a potential membrane glycoprotein. *Virology* **165,** 151–164.

Kouzarides, T., Bankier, A. T., Satchwell, S. C., Weston, K., Tomlinson, P., and Barrell, B. G. (1987). Large-scale rearrangement of homologous regions in the genomes of HCMV and EBV. *Virology* **157,** 397–413.

Lawrence, G. L., Chee, M., Craxton, M. A., Gompels, U. A., Honess, R. W., and Barrell, B. G. (1990). Human herpesvirus 6 is closely related to human cytomegalovirus. *J. Virol.* **64,** 287–299.

Lawrence, G. L., Nicholas, J., and Barrell, B. G. (1995). Human herpesvirus 6 (strain U1102) encodes homologues of the conserved herpesvirus glycoprotein gM and the alphaherpesvirus origin-binding protein. *J. Gen. Virol.* **76,** 147–152.

Leach, C. T., Newton, E. R., McParlin, S., and Jenson, H. B. (1994). Human herpesvirus 6 infection of the female genital tract. *J. Infect. Dis.* **169,** 1281–1283.

Levy, J. A., Landay, A., and Lennette, E. T. (1990). Human herpesvirus 6 inhibits human immunodeficiency virus type 1 replication in cell culture. *J. Clin. Microbiol.* **28,** 2362–2364.

Lindquester, G. J., and Pellett, P. E. (1991). Properties of the human herpesvirus 6 strain Z29 genome: G + C content, length, and presence of variable-length directly repeated terminal sequence elements. *Virology* **182,** 102–110.

Littler, E., Lawrence, G., Liu, M. Y., Barrell, B. G., and Arrand, J. R. (1990). Identification, cloning, and expression of the major capsid protein gene of human herpesvirus 6. *J. Virol.* **64,** 714–722.

Littler, E., Stuart, A. D., and Chee, M. S. (1992). Human cytomegalovirus UL97 open reading frame encodes a protein that phosphorylates the antiviral nucleoside analogue ganciclovir. *Nature (London)* **358,** 160–162.

Liu, B., and Stinski, M. F. (1992). Human cytomegalovirus contains a tegument protein that enhances transcription from promoters with upstream ATF and AP-1 cis-acting elements. *J. Virol.* **66,** 4434–4444.

Liu, D. X., Gompels, U. A., Foa-Tomasi, L., and Campadelli-Fiume, G. (1993a). Human herpesvirus-6 glycoprotein H and L homologs are components of the gp100 complex and the gH external domain is the target for neutralizing monoclonal antibodies. *Virology* **197,** 12–22.

Liu, D. X., Gompels, U. A., Nicholas, J., and Lelliott, C. (1993b). Identification and expression of the human herpesvirus 6 glycoprotein H and interaction with an accessory 40K glycoprotein. *J. Gen. Virol.* **74,** 1847–1857.

Liu, F., and Roizman, B. (1991). The herpes simplex virus 1 gene encoding a protease also contains within its coding domain the gene encoding the more abundant substrate. *J. Virol.* **65,** 5149–5156.

Locker, H., and Frenkel, N. (1979). BamI, KpnI, and SalI restriction enzyme maps of the DNAs of herpes simplex virus strains Justin and F: Occurrence of heterogeneities in defined regions of the viral DNA. *J. Virol.* **32,** 429–441.

Locker, H., Frenkel, N., and Halliburton, I. (1982). Structure and expression of class II defective herpes simplex virus genomes encoding infected cell polypeptide number 8. *J. Virol.* **43,** 574–593.

Lopez, C., Pellett, P., Stewart, J., Goldsmith, C., Sandelin, K., Black, J., Warfield, D., and Feorino, P. (1988). Characteristics of human herpesvirus-6. *J. Infect. Dis.* **157,** 1271–1273.

Lusso, P., De, M. A., Malnati, M., Lori, F., DeRocco, S. E., Baseler, M., and Gallo, R. C. (1991). Induction of CD4 and susceptibility to HIV-1 infection in human CD8+ T lymphocytes by human herpesvirus 6. *Nature (London)* **349,** 533–535.

Lusso, P., Ensoli, B., Markham, P. D., Ablashi, D. V., Salahuddin, S. Z., Tschachler, E., WongStaal, F., and Gallo, R. C. (1989). Productive dual infection of human CD4+ T lymphocytes by HIV-1 and HHV-6. *Nature (London)* **337,** 370–373.

Lusso, P., Malnati, M. S., Garzino-Demo, A., Crowley, R. W., Long, E. O., and Gallo, R. C. (1993). Infection of natural killer cells by human herpesvirus 6. *Nature (London)* **362,** 458–462.

Lusso, P., Secchiero, P., Crowley, R. W., Garzino-Demo, A., Berneman, Z. N., and Gallo, R. C. (1994). CD4 is a critical component of the receptor for human herpesvirus 7: Interference with human immunodeficiency virus. *Proc. Natl. Acad. Sci. USA* **91,** 3872–3876.

Martin, M., Thomson, B. J., Honess, R. W., Craxton, M. A., Gompels, U. A., Liu, M. Y., Littler, E., Arrand, J. R., Teo, I., and Jones, M. D. (1991a). The genome of human herpesvirus 6: Maps of unit-length and concatemeric genomes for nine restriction endonucleases. *J. Gen. Virol.* **72,** 157–168.

Martin, M. E. D., Nicholas, J., Thomson, B. J., Newman, C., and Honess, R. W. (1991b). Identification of a transactivating function mapping to the putative immediate-early locus of human herpesvirus 6. *J. Virol.* **65,** 5381–5390.

Masse, M. J. O., Karlin, S., Schachtel, G., and Mocarski, E. S. (1992). Human cytomegalovirus origin of DNA replication (oriLyt) resides within a highly complex repetitive region. *Proc. Natl. Acad. Sci. USA* **89,** 5246–5250.

McGeoch, D. J. (1990). Evolutionary relationships of virion glycoprotein genes in the S regions of alphaherpesvirus genomes. *J. Gen. Virol.* **71,** 2361–2367.

McGeoch, D. J. (1990). Protein sequence comparisons show that the 'pseudoproteases' encoded by poxviruses and certain retroviruses belong to the deoxyuridine triphosphatase family. *Nucleic Acids Res.* **18,** 4105–4110.

McGeoch, D. J., and Cook, S. (1994). Molecular phylogeny of the alphaherpesvirinae subfamily and a proposed evolutionary timescale. *J. Mol. Biol.* **238,** 9–22.

McGeoch, D. J., Cunningham, C., McIntyre, G., and Dolan, A. (1991). Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2. *J. Gen. Virol.* **72,** 3057–3075.

McGeoch, D. J., Dalrymple, M. A., Davison, A. J., Dolan, A., Frame, M. C., McNab, D., Perry, L. J., Scott, J. E., and Taylor, P. (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.

McLauchlan, J., Phelan, A., Loney, C., Sandri-Goldin, R. M., and Clements, J. B. (1992). Herpes simplex virus IE63 acts at the posttranscriptional level to stimulate viral mRNA 3′ processing. *J. Virol.* **66,** 6939–6945.

Messerle, M., Keil, G. M., and Koszinowski, U. H. (1991). Structure and expression of murine cytomegalovirus immediate-early gene 2. *J. Virol.* **65,** 1638–1643.

Minton, E. J., Tysoe, C., Sinclair, J. H., and Sissons, J. (1994). Human

cytomegalovirus infection of the monocyte/macrophage lineage in bone marrow. *J. Virol.* **68,** 4017–4021.

Mocarski, E., and Roizman, B. (1982). Structure and role of the herpes simplex virus DNA termini in inversion, circularization and generation of virion DNA. *Cell* **31,** 89–97.

Mocarski, E. S., Deiss, L. P., and Frenkel, N. (1985). Nucleotide sequence and structural features of a novel U(S)-a junction present in a defective herpes simplex virus genome. *J. Virol.* **55,** 140–146.

Neipel, F., Ellinger, K., and Fleckenstein, B. (1991). The unique region of the human herpesvirus 6 genome is essentially collinear with the U(L) segment of human cytomegalovirus. *J. Gen. Virol.* **72,** 2293–2297.

Neipel, F., Ellinger, K., and Fleckenstein, B. (1992). Gene for the major antigenic structural protein (p100) of human herpesvirus 6. *J. Virol.* **66,** 3918–3924.

Neote, K., DiGregorio, D., Mak, J. Y., Horuk, R., and Schall, T. J. (1993). Molecular cloning, functional expression and signaling characteristics of a c-c chemokine receptor. *Cell* **72,** 415–425.

Newcomb, W. W., Homa, F. L., Thomsen, D. R., Ye, Z., and Brown, J. C. (1994). Cell-free assembly of the herpes simplex virus capsid. *J. Virol.* **68,** 6059–6063.

Nicholas, J. (1994). Nucleotide sequence of a 21-kilobase-pair region of the genome of human herpesvirus-6 containing homologues of human cytomegalovirus major immediate-early and replication genes. *Virology* **204,** 738–750.

Nicholas, J., Cameron, K. R., and Honess, R. W. (1992). Herpesvirus saimiri encodes homologues of G protein-coupled receptors and cyclins. *Nature (London)* **355,** 362–365.

Nicholas, J., and Martin, M. (1994). Nucleotide sequence analysis of a 38.5-kilobase-pair region of the genome of human herpesvirus 6 encoding human cytomegalovirus immediate-early gene homologs and transactivating functions. *J. Virol.* **68,** 597–610.

O'Brien, S. J., Seuanez, H. N., and Womack, J. E. (1988). Mammalian genome organisation: an evolutionary view. *Annu. Rev. Genet.* **22,** 232–251.

Okuno, T., Takahashi, K., Balachandra, K., Shiraki, K., Yamanishi, K., Takahashi, M., and Baba, K. (1989). Seroepidemiology of human herpesvirus 6 infection in normal children and adults. *J. Clin. Microbiol.* **27,** 651–653.

Ong, C. K., Chan, S. Y., Campo, M. S., Fujinaga, K., Mavromara-Nazos, P., Labropoulou, V., Pfister, H., Tay, S. K., Ter, M. J., Villa, L. L., and Bernard, H. U. (1993). Evolution of human papillomavirus type 18: An ancient phylogenetic root in Africa and intratype diversity reflect coevolution with human ethnic groups. *J. Virol.* **67,** 6424–6431.

Oppenheim, J. J., Zachariae, C. O. C., Mukaida, N., and Matsushima, K. (1991). Properties of the novel proinflammatory supergene "intercrine" cytokine family. *Annu. Rev. Immunol.* **9,** 617–648.

Packham, G., Brimmell, M., Cook, D., Sinclair, A. J., and Farrell, P. J. (1993). Strain variation in Epstein-Barr virus immediate early genes. *Virology* **192,** 541–550.

Pari, G. S., and Anders, D. G. (1993). Eleven loci encoding trans-acting factors are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA replication. *J. Virol.* **67,** 6979–6988.

Parker, B. D., Bankier, A., Satchwell, S., Barrell, B., and Farrell, H. (1990). Sequence and transcription of Raji Epstein-Barr virus DNA spanning the B95-8 deletion region. *Virology* **179,** 339–346.

Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85,** 2444–2448.

Pellett, P. E., Sanchez-Martinez, D., Dominguez, G., Black, J. B., Anton, E., Greenamoyer, C., and Dambaugh, T. R. (1993). A strongly immunoreactive virion protein of human herpesvirus 6 variant B strain Z29: Identification and characterization of the gene and mapping of a variant-specific monoclonal antibody reactive epitope. *Virology* **195,** 521–531.

Pfeiffer, B., Berneman, Z. N., Neipel, F., Chang, C. K., Tirwatnapong, S., and Chandran, B. (1993). Identification and mapping of the gene encoding the glycoprotein complex gp82-gp105 of human herpesvirus 6 and mapping of the neutralizing epitope recognized by monoclonal antibodies. *J. Virol.* **67,** 4611–4620.

Phelan, A., CarmoFonseca, M., McLauchlan, J., Lamond, A. I., and Clements, J. B. (1993). A herpes simplex virus type 1 immediate-early gene product, IE63, regulates small nuclear ribonucleoprotein distribution. *Proc. Natl. Acad. Sci. USA* **90,** 9056–9060.

Pilling, A., Davison, A. J., Telford, E., and Meredith, D. M. (1994). The equine herpesvirus type 1 glycoprotein homologous to herpes simplex virus type 1 glycoprotein M is a major constituent of the virus particle. *J. Gen. Virol.* **75,** 439–442.

Poon, A., and Roizman, B. (1993). Characterization of a temperature-sensitive mutant of the UL15 open reading frame of herpes simplex virus 1. *J. Virol.* **67,** 4497–4503.

Preston, V. G., AlKobaisi, M. F., McDougall, I. M., and Rixon, F. J. (1994). The herpes simplex virus gene UL26 proteinase in the presence of the UL26.5 gene product promotes the formation of scaffold-like structures. *J. Gen. Virol.* **75,** 2355–2366.

Prezioso, P. J., Cangiarella, J., Lee, M., Nuovo, G. J., Borkowsky, W., Orlow, S. J., and Greco, M. A. (1992). Fatal disseminated infection with human herpesvirus-6. *J. Pediatrics* **120,** 921–923.

Qian, G., Wood, C., and Chandran, B. (1993). Identification and characterization of glycoprotein gH of human herpesvirus-6. *Virology* **194,** 380–386.

Roizman, B., Desrosiers, R. C., Fleckenstein, B., Lopez, C., Minson, A. C., and Studdert, M. J. (1992). The family herpesviridae: An update. *Arch. Virol.* **123,** 425–449.

Sakaoka, H., Saito, H., Sekine, K., Aomori, T., Grillner, L., Wadell, G., and Fujinaga, K. (1987). Genomic comparison of herpes simplex virus type 1 isolates from Japan, Sweden and Kenya. *J. Gen. Virol.* **68,** 749–764.

Salahuddin, S. Z., Ablashi, C. V., Markham, P. D., Josephs, S. F., Sturzenegger, S., Kaplan, M., Halligan, G., Biberfield, P., Wong-Staal, F., Kramarsky, B., and Gallo, R. C. (1986). Isolation of a new virus, HBLV, in patients with lymphoproliferative disorders. *Science* **234,** 596–601.

Sample, J., Young, L., Martin, B., Chatman, T., Kieff, E., Rickinson, A., and Kieff, E. (1990). Epstein–Barr virus types 1 and 2 differ in their EBNA-3A, EBNA-3B, and EBNA-3C genes. *J. Virol.* **64,** 4084–4092.

Sandford, G. R., Ho, K., and Burns, W. H. (1993). Characterization of the major locus of immediate-early genes of rat cytomegalovirus. *J. Virol.* **67,** 4093–4103.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.

Sankoff, D., Cedergreen, R., and Abel, Y. (1990). Genomic divergence through gene rearrangement. *Methods Enzymol.* **183,** 428–438.

Schiewe, U., Neipel, F., Schreiner, D., and Fleckenstein, B. (1994). Structure and transcription of an immediate-early region in the human herpesvirus 6 genome. *J. Virol.* **68,** 2978–2985.

Staden, R. (1987). Computer handling of DNA sequencing projects. *In* "Nucleic Acid and Protein Sequence Analysis: A Practical Approach" (M. J. Bishop and C. J. Rawlings, Eds.) IRL Press, Oxford.

Stasiak, P. C., and Mocarski, E. S. (1992). Transactivation of the cytomegalovirus ICP36 gene promoter requires the alpha gene product TRS1 in addition to IE1 and IE2. *J. Virol.* **66,** 1050–1058.

Sullivan, V., Biron, K. K., Talarico, C., Stanat, S. C., Davis, M., Pozzi, L. M., and Coen, D. M. (1993). A point mutation in the human cytomegalovirus DNA polymerase gene confers resistance to ganciclovir and phosphonylmethoxyalkyl derivatives. *Antimicrob. Agents Chemother.* **37,** 19–25.

Sullivan, V., Talarico, C. L., Stanat, S. C., Davis, M., Coen, D. M., and Biron, K. K. (1992). A protein kinase homologue controls phosphorylation of ganciclovir in human cytomegalovirus-infected cells. *Nature (London)* **358,** 162–164.

Suzuki, M., Savoysky, E., Izuta, S., Tatebe, M., Okajima, T., and Yoshida, S. (1993). RNA priming coupled with DNA synthesis on natural template by calf thymus DNA polymerase alpha-primase. *Biochemistry (USA)* **32,** 12782–12792.

Takahashi, K., Sonoda, S., Higashi, K., Kondo, T., Takahashi, H., Takahashi, M., and Yamanishi, K. (1989). Predominant CD4 T-lymphocyte tropism of human herpesvirus 6-related virus. *J. Virol.* **63,** 3161–3163.

Tanaka, K., Kondo, T., Torigoe, S., Okada, S., Mukai, T., and Yamanishi, K. (1994). Human herpesvirus 7: Another causal agent for roseola (exanthem subitum). *J. Pediatr.* **125,** 1–5.

Taylor-Wiedeman, J., Sissons, J., Borysiewicz, L. K., and Sinclair, J. H. (1991). Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. *J. Gen. Virol.* **72,** 2059–2064.

Taylor-Wiedeman, J., Sissons, P., and Sinclair, J. (1994). Induction of endogenous human cytomegalovirus gene expression after differentiation of monocytes from healthy carriers. *J. Virol.* **68,** 1597–1604.

Tedder, T. S., Briggs, M., Cameron, C. H., Honess, R. W., Robertson, D., and White, H. (1987). A novel lymphotropic herpesvirus. *Lancet* **2,** 390–391.

Telford, E., Studdert, M. J., Agius, C. T., Watson, M. S., Aird, H. C., and Davison, A. J. (1993). Equine herpesviruses 2 and 5 are gammaherpesviruses. *Virology* **195,** 492–499.

Telford, E., Watson, M. S., McBride, K., and Davison, A. J. (1992). The DNA sequence of equine herpesvirus-1. *Virology* **189,** 304–316.

Teo, I. A., Griffin, B. E., and Jones, M. D. (1991). Characterization of the DNA polymerase gene of human herpesvirus 6. *J. Virol.* **65,** 4670–4680.

Thompson, J., Choudhury, S., Kashanchi, F., Doniger, J., Berneman, Z., Frenkel, N., and Rosenthal, L. J. (1994). A transforming fragment within the direct repeat region of human herpesvirus type 6 that transactivates HIV-1. *Oncogene* **9,** 1167–1175.

Thompson, R., Honess, R. W., Taylor, L., Morran, J., and Davison, A. J. (1987). Varicella-zoster virus specifies a thymidylate synthase. *J. Gen. Virol.* **68,** 1449–1455.

Thomsen, D. R., Roof, L. L., and Homa, F. L. (1994). Assembly of herpes simplex virus (HSV) intermediate capsids in insect cells infected with recombinant baculoviruses expressing HSV capsid proteins. *J. Virol.* **68,** 2442–2457.

Thomson, B. J., Dewhurst, S., and Gray, D. (1994a). Structure and heterogeneity of the a sequences of human herpesvirus 6 strain variants U1102 and Z29 and identification of human telomeric repeat sequences at the genomic termini. *J. Virol.* **68,** 3007–3014.

Thomson, B. J., Efstathiou, S., and Honess, R. W. (1991). Acquisition of the human adeno-associated virus type-2 rep gene by human herpesvirus type-6. *Nature (London)* **351,** 78–80.

Thomson, B. J., and Honess, R. W. (1992). The right end of the unique region of the genome of human herpesvirus 6 U1102 contains a candidate immediate early gene enhancer and a homologue of the human cytomegalovirus US22 gene family. *J. Gen. Virol.* **73,** 1649–1660.

Thomson, B. J., Weindler, F. W., Gray, D., Schwaab, V., and Heilbronn, R. (1994b). Human herpesvirus 6 (HHV-6) is a helper virus for adeno-associated virus type 2 (AAV-2) and the AAV-2 rep gene homologue in HHV-6 can mediate AAV-2 DNA replication and regulate gene expression. *Virology* **204,** 304–311.

Von Heijne, G., and Gavel, Y. (1988). Topogenic signals in integral membrane proteins. *Eur. J. Biochem. FEBS* **174,** 671–678.

Waga, S., and Stillman, B. (1994). Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication *in vitro*. *Nature (London)* **369,** 207–212.

Ward, K. N., and Gray, J. J. (1994). Primary human herpesvirus-6 infection is frequently overlooked as a cause of febrile fits in young children. *J. Med. Virol.* **42,** 119–123.

Ward, K. N., Gray, J. J., Fotheringham, M. W., and Sheldon, M. J. (1993). IgG antibodies to human herpesvirus-6 in young children: Changes in avidity of antibody correlate with time after infection. *J. Med. Virol.* **39,** 131–138.

Welch, A. R., McNally, L. M., Hall, M. R. T., and Gibson, W. (1993). Herpesvirus proteinase: Site-directed mutagenesis used to study maturational, release and inactivation cleavage sites of precursor and to identify a possible catalytic site serine and histidine. *J. Virol.* **67,** 7360–7372.

Wilborn, F., Brinkmann, V., Schmidt, C. A., Neipel, F., Gelderblom, H., and Siegert, W. (1994). Herpesvirus type 6 in patients undergoing bone marrow transplantation: Serologic features and detection by polymerase chain reaction. *Blood* **83,** 3052–3058.

Williams, A. F., and Barclay, A. N. (1988). The immunoglobulin superfamily-domains for cell surface recognition. *Annu. Rev. Immunol.* **6,** 381–405.

Winkler, M., Rice, S. A., and Stamminger, T. (1994). UL69 of human cytomegalovirus, an open reading frame with homology to ICP27 of herpes simplex virus, encodes a transactivator of gene expression. *J. Virol.* **68,** 3943–3954.

Yamanishi, K., Okuno, T., Shiraki, K., Takahashi, M., Kondo, T., Asano, Y., and Kurata, T. (1988). Identification of human herpesvirus-6 as a causal agent for exanthem subitum. *Lancet* **1,** 1065–1067.

Yang, C. L., Chang, L. S., Zhang, P., Hao, H., Zhu, L., Toomey, N. L., and Lee, M. Y. (1992). Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase delta. *Nucleic Acids Res.* **20,** 735–745.

Zhou, Y., Chang, C. K., Qian, G., Chandran, B., and Wood, C. (1994). trans-Activation of the HIV promoter by a cDNA and its genomic clones of human herpesvirus-6. *Virology* **199,** 311–322.