Available online at www.sciencedirect.com

SCIENCE DIRECT

Virology 331 (2005) 429–440

VIROLOGY

www.elsevier.com/locate/yviro

Complete genome sequence of cercopithecine herpesvirus 2 (SA8) and comparison with other simplexviruses

Shaun D. Tyler, Geoffrey A. Peters, Alberto Severini*

National Microbiology Laboratory, Canadian Science Centre for Human and Animal Health, 1015 Arlington Street, Winnipeg, Manitoba, Canada, R3E 3R2

Received 17 August 2004; returned to author for revision 8 September 2004; accepted 24 September 2004
Available online 13 November 2004

Abstract

We have obtained the complete sequence of the herpesvirus simian agent 8 (SA8; cercopithecine herpesvirus 2) a baboon simplexvirus closely related to the monkey B virus and herpes simplex virus types 1 and 2. The genome of SA8 is 150,715 bp long, with an overall G/C content of 76%, the highest among the simplexviruses sequenced so far. The sequencing has confirmed that the genomic arrangement of SA8 is similar to that of other simplexviruses: unique long and unique short regions bordered by two sets of inverted repeats. All genes identified in SA8 are homologous and collinear with those of the monkey B virus, including the lack of the RL1 open reading frame, a gene responsible for neurovirulence in human herpes simplex viruses. This latter finding supports the hypothesis that a different pathogenetic mechanism may have developed in human simplexviruses, after their divergence from monkey simplexviruses.

Keywords: Herpesvirus; Complete genome; Genomics; Simplexvirus; SA8; Cercopithecine herpesvirus 2

Introduction

Simian agent 8 (SA8; cercopithecine herpesvirus 2) is an alpha-herpesvirus (family: *Herpesviridae*; subfamily: *Herpesvirinae*) that was initially isolated from an African green monkey in 1958 and classified as a herpesvirus on the basis of its characteristics in cell culture and neurotropism in monkeys and in experimentally inoculated rabbits (Malherbe and Harwin, 1958). Subsequent isolation and serological studies indicated that the natural hosts were African monkeys of the genus Papio (baboons) (Malherbe and Strickland-Cholmley, 1969a,b; Veit et al., 1996).

SA8 shares marked biological (Borchers and Ozel, 1993; Brack et al., 1985; Levin et al., 1988), antigenic (Eberle et al., 1989; Borchers et al., 1990; Katz et al., 2002), and sequence homologies (Eberle and Black, 1991, 1993; Eberle et al., 1993) with other alpha-herpesviruses, and it has been included in the genus simplexvirus, together with the herpes B virus (cercopithecine herpesvirus 1), herpes virus papio type 2 (cercopithecine herpesvirus 16) (Eberle et al., 1995), and herpes simplex virus types 1 and 2 (human herpesviruses 1 and 2) (van Regenmortel et al., 2000). Infections with SA8 in baboons cause a disease that is very similar to genital herpes in humans, produced by HSV-2. Like HSV-2, SA8 is primarily sexually transmitted and in most cases the infection is asymptomatic. When symptomatic, SA8 causes recurrent lesions of the genital and oral mucosae (Levin et al., 1988; Martino et al., 1998) and it is more severe in females than males (Martino et al., 1998). Severe infections of newborn monkeys resembling human neonatal herpes have also been documented (Brack et al., 1985; Eichberg et al., 1976). Evidence shows that SA8, like other primate alpha-herpesviruses, establishes latency in the sensory ganglia (Kalter et al., 1978). SA8 is not known to cause diseases in humans, but it can be a source of considerable morbidity in breeding colonies of baboons (Levin et al., 1988).

Partial sequencing of the SA8 genome has shown that gene products share extensive homology with HSV-1 and

* Corresponding author. Fax: +1 204 789 2140.
E-mail address: Alberto_Severini@hc-sc.gc.ca (A. Severini).

0042-6822/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2004.09.042

HSV-2 (Eberle et al., 1993). It has also been shown that the overall genome structure follows the "type E arrangement" (Roizman and Pellet, 2001) of other simplexviruses, with unique long (UL) and unique short (US) sequences flanked by two sets of long and short inverted repeats (Eberle et al., 1993). The genes of the short region have been sequenced and they were all shown to be conserved and collinear with the other simplexviruses (Eberle et al., 1993), in particular to the herpes B virus, a primate alpha-herpesvirus that infects Asian macaques. The entire genome of the B virus has recently been sequenced (Perelygina et al., 2003) and it was shown to be collinear and homologous to the genomes of HSV-1 and HSV-2, with two important differences: (i) B virus lacks a homologue of the HSV open reading frame γ34.5 (RL1), a gene responsible for neurovirulence in mice (Whitley et al., 1993) and that inhibits protein synthesis shutoff by the host RNA-dependent protein kinase (Chou et al., 1995); (ii) B virus has two copies of each of the three origins of replication found in HSV-1 and HSV-2.

While accidental transmission of the B virus to humans or other African primates results in severe infections and devastating encephalomyelitis (Davidson and Hummeler, 1960; Sabin, 1949; Sabin and Wright, 1934; Weigler, 1992), SA8 is not known to cause diseases in primates outside the natural hosts. This is an interesting difference, and comparing the genome and gene products of SA8 with those of other simplexviruses may provide clues for the mechanism of pathogenesis and neuroinvasiveness of simplexviruses. We have therefore produced the complete sequence of SA8 (strain B264) (Malherbe and Harwin, 1963), and we have performed a comparative analysis with the genomes of other simplexviruses. We have identified all the open reading frames and determined the sequences of the genomic termini and origins of replication.

Results and discussion

Overall genomic organization

The genome of SA8 was found to be 150,715 bp long (GenBank accession number AY714813), a size similar to HSV-1 (152 kb), HSV-2 (155 kb), and B virus (157 kb), and like for these viruses it was composed of a unique long (UL) and unique short (US) region delimited by long (RL) and short (RS) inverted repeats, as was previously shown (Eberle et al., 1993). This arrangement, defined as type E, is typical of simplexviruses and other alpha- and beta-herpesviruses (Roizman and Pellet, 2001). Homologous recombination events at the inverted repeats during replication lead to the generation of four isomers that differ in the mutual orientation of the UL and US regions. The genome arrangement in SA8 is shown in Fig. 1, in the "prototype" configuration. The overall G/C content of the genome was 76%, highest among the primate simplexviruses: 74.5% for B virus, 68.1% for HSV-1, and 70.4% for HSV-2. The inverted repeat regions had the highest G/C content (82.1% in the long repeat and 79.7% for the short repeat), but the G/C content was high throughout the genome and the long and short unique regions had a G/C content of 74.2% and 76.1%, respectively. The coding regions had an overall G/C content of 74.6%.

The overall homology at the DNA level was 83.3% between SA8 and B virus (GenBank accession number NC_004812) (Perelygina et al., 2003), 64.1% and 68.8% between SA8 and HSV-1 (NC_001806) (Dolan et al., 1992, 1998; McGeoch et al., 1985, 1986, 1988; Perry and McGeoch, 1988), and HSV-2 (NC_001798) (Barnett et al., 1992; Dolan et al., 1998; Everett and Fenwick, 1990; McGeoch et al., 1987, 1991), respectively.

As with a number of other herpesviruses, the genome of SA8 contained numerous sets of short (2–63 bp) tandemly repeated sequences. Forty-nine sets of perfect repeats were present and an additional 77 sets of imperfect repeats, in which one or more unit diverged slightly from the consensus sequence, were also identified (data not shown). For comparison, we performed a similar analysis on the published sequences of HSV-1, HSV-2, and B virus and we detected a total of 39, 49, and 76 sets, respectively, including perfect and imperfect repeats.

Origins of replication

Three putative origins of replication were detected in the SA8 genome, one in the unique long region (oriL) and two in the inverted repeat regions (oriS). The origins sequences were not tandemly repeated as in the B virus (Perelygina et al., 2003), but instead the three putative origins contained an identical inverted repeat 96 bp in length. This is unlike the situation in HSV-1 and HSV-2 where the sequence of oriS has diverged somewhat from the consensus of the other origins of primate simplexviruses. As shown in Fig. 2, the origin sequences of SA8 show a close homology with the origins of the monkey B virus and with the oriL origins of HSV-1 and HSV-2. Fig. 2 also shows that the oriS in HSV-1 and HSV-2 diverge in the sequence of box III′, and it remains to be determined if the lack of box III′ results in a different modality of usage of these origins of replication.

The 'a' sequence and genomic termini

Herpesvirus genomes are linear and they carry at both ends directly repeated 'a' sequences that contain the *cis*-acting signals for cleavage and packaging of the concatemeric genomes. The cleavage/packaging signals consist of two conserved regions called pac1 and pac2 (McVoy et al., 1998; McVoy et al., 2000; Smiley et al., 1990; Varmuza and Smiley, 1985). One or more copies of the 'a' sequence in inverted orientation are also found at the junction between the internal long and short repeat regions of the genomes of simplexviruses (Mocarski and Roizman, 1982; Perelygina et al., 2003). The internal 'a' sequence of SA8 (shown in

Fig. 3 in direct orientation) was bordered by a set of direct repeats, here defined DR1 in keeping with the nomenclature adopted for the ‘a’ sequence of HSV, which is also bordered by direct repeats (Mocarski and Roizman, 1982). There was a second set of tandemly repeated sequences, indicated as DR2 in Fig. 3. Fig. 3 also highlights the packaging signals pac1 and pac2, which are conserved in B virus (Perelygina et al., 2003), HSV-1, HSV-2, and other herpesviruses (Deiss et al., 1986; McVoy et al., 1998). Two copies of a variant (CGGGGCG) of the pac2 CGCGGCG motif, described for HSV-1, HSV-2, and B virus, are found at the expected position, downstream of the pac2 A-rich region. This variation of the motif is also found in varicella–zoster virus (Davison, 1984), and therefore we postulate that it may serve the same role as the classic motif in the other simplexviruses. Three copies of the CGCGGCG motif are found elsewhere in the DR1 repeats and in the ‘a’ sequence of SA8, as shown in Fig. 3; however, their presence in such G/C rich regions may be coincidental.

In order to identify the sequences of the genomic termini, the approximate location of the ‘a’ sequence was determined by comparison of the junction between the internal repeat regions with the ‘a’ sequences from HSV-1 and HSV-2 and B virus. Computer-based mapping was then used to identify restriction enzymes that would facilitate the selective isolation of the terminal ‘a’ sequences. Left terminus clones contained a complete ‘a’ sequence with both DR1 repeats and terminated at the last nucleotide of the left DR1 repeat. Right terminus clones contained most ‘a’ sequence but ended one nucleotide before the right DR1 sequence (see Fig. 3).

Joining together the L and S termini would reconstitute an ‘a’ sequence in which a single nucleotide just upstream of DR1 is missing. However, our method for blunt-ending the genomic DNA before cloning (see Materials and methods) uses a combination of T4 polymerase and Klenow polymerase. Because of the very strong 3′→5′ exonuclease activity of the T4 polymerase, fragments with a 3′ overhang will be blunt ended by removal of one nucleotide. This means that our result is consistent with a cleavage mechanism of concatemers that leaves a 1-bp 3′ overhang (a C on the L terminus and a G on the S terminus). This mechanism is analogous to what was described for cleavage of HSV-1 concatemers, in which evidence indicated a similar G/C 3′ overhang at the genomic termini (Mocarski and Roizman, 1982). According to a proposed model for HSV replication, circularization of the herpesvirus genome would occur by end-to-end ligation soon after entry in the host cell. End-to-end ligation of the SA8 genome would create a junction with two ‘a’ sequences that share one DR1 repeat, as previously described for HSV-1 (Mocarski and Roizman, 1982). Alternative models suggest that circular-

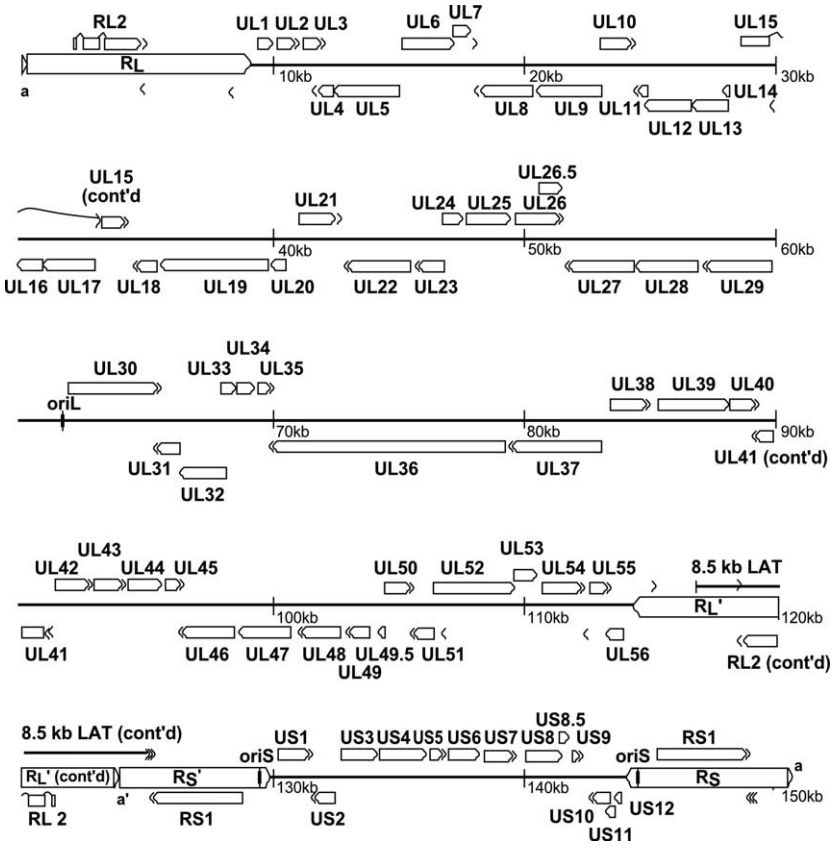

Fig. 1. Map of the open reading frames and main features of the SA8 genome. The positions of polyadenylation sites are indicated by single chevrons pointing in the predicted direction of transcription.

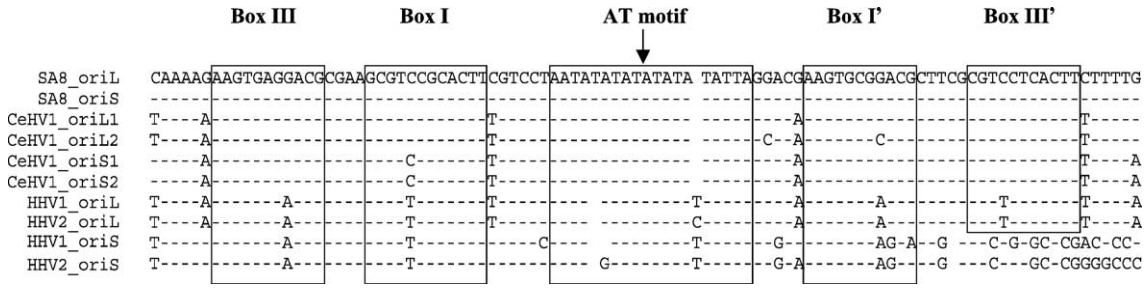

Fig. 2. Sequences of the inverted repeat region of the SA8 origins of replication and comparison with the origins of replication of the B virus, HSV-1, and HSV-2. The arrow indicates the center of the inverted repeat. The boxes indicate the sites that bind the origin binding protein in HSV-1, as previously defined (Hazuda et al., 1991).

ization would occur by homologous recombination of the terminal 'a' sequences (Roizman, 1979; Yao et al., 1997) or that replication of the linear genome occurs at the same time as homologous recombination (Jackson and DeLuca, 2003), which would produce junctions with a single 'a' sequence bordered by two DR1 repeats.

All of the clones that we sequenced contained a single copy of the 'a' sequence, in contrast to what was previously reported for HSV-1 (Mocarski and Roizman, 1982) and the B virus (Perelygina et al., 2003), in which multiple copies of the 'a' sequence were found at the L terminus and at the internal junction between the long and short repeat regions, but not at the S terminus. We therefore performed a Southern blot of intracellular SA8 DNA (containing both replicating and mature genomes), digested with restriction enzymes, and hybridized with a probe that spans the internal junction between L and S repeats (see map in Fig. 4B). Fragments of the predicted sizes were detected corresponding to the short and long terminal sequences as well as the internal junction (Fig. 4A). Additional bands were also detected that are of the correct size for what would be predicted if multiple 'a' sequences (236 bp) were present on the long terminal fragment and at the internal junction. At least three additional bands were clearly seen for the internal junction and two for the long terminal sequence (it should be noted that if more than two additional 'a' sequences were present on the long terminus the resulting sizes would comigrate with the internal junction fragments and would not appear as distinct bands). No additional fragments were seen to indicate more than one copy of the 'a' sequence on the short terminal fragment.

Open reading frames

Open reading frames were initially identified by homology analysis of predicted gene products with the published sequences of HSV-1, HSV-2, and B virus. Previous partial sequencing of the short unique region had shown that all the predicted open reading frames of SA8 were homologous and collinear with their counterparts in other simplexviruses. We have confirmed this finding and our data show that all predicted open reading frames in the long unique region are also collinear and homologous to those in B virus, HSV-1, and HSV-2 (Table 1). Two genes, RL2 and UL15, had introns and their organization matched the ones in B virus and HSV. As in B virus (Perelygina et al., 2003), the RL1

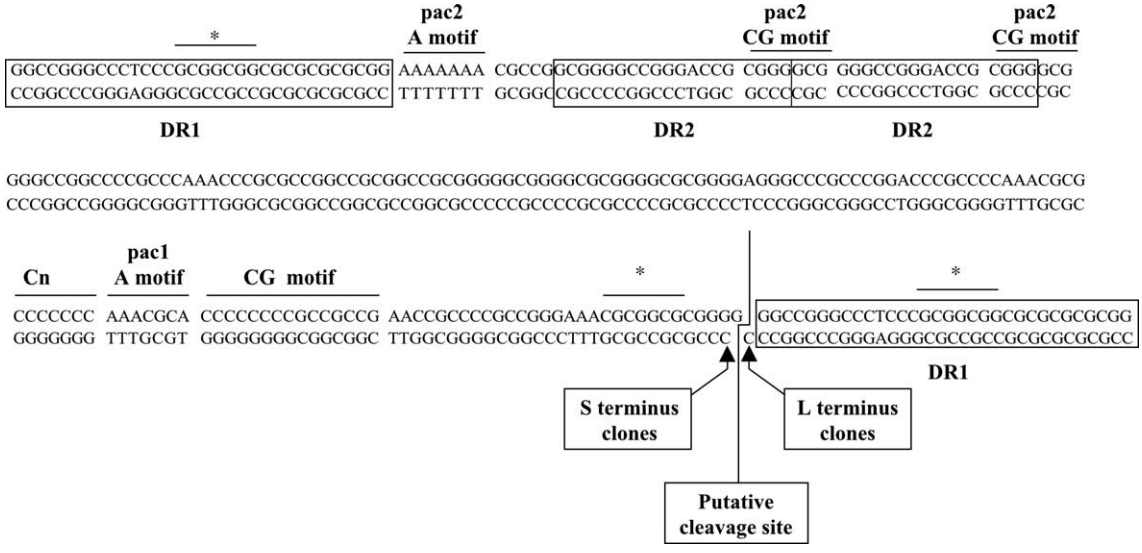

Fig. 3. Sequence and features of the internal 'a' region of SA8. The boxes labeled DR1 highlight the direct terminal repeats found at both ends of the internal 'a' sequence. The thin lines with the asterisks indicate occurrences of the GCGGCGG motif that is found in the pac2 site of the B virus and HSV (see text).

open reading frame (neurovirulence gene, γ34.5) was not detected. The region of the SA8 genome corresponding to RL1 in HSV was translated into the six possible frames, and the predicted products were examined for homology with the RL1 gene of HSV-1 or with other protein sequences in the NCBI databases. No significant homology with other proteins or motifs was detected, strengthening the conclusion that open reading frames are not present in this region of simian simplexviruses. The RL1 gene product has been shown to have two distinct pathogenetic functions: it is a neurovirulence factor in mice (Chou et al., 1990) and an inhibitor of the interferon-induced protein kinase R system (Chou et al., 1995). The homology of RL1 protein with the conserved mammalian protein GADD34 family of vertebrate proteins (Chou and Roizman, 1994; Fornace et al., 1989; Lord et al., 1990; McGeoch and Barnett, 1991; Zhan et al., 1994) and its absence from other alpha-herpesviruses suggests that this gene may have been acquired during the evolution of simplexviruses. The lack of RL1 in the closely related B virus and SA8 suggests that the acquisition of this pathogenetic mechanism may have occurred in the common progenitor of the human simplexvirus, soon after their divergence from the cercopithecine herpesviruses. Following this hypothesis, the ancestral mode of inhibition of the PRK-mediated protein synthesis shut-off could be based on the US11 protein, which in HSV-1 has been shown to be expressed as a result of mutations that compensate for the deletion of the RL1 gene (Cassady et al., 1998; Mohr and Gluzman, 1996; Mulvey et al., 1999, 2004). More support for this hypothesis comes from recent evidence that the B virus US11 protein acts as an inhibitor of the PRK system in vitro (Zhu et al., 2004). It is however possible that the acquisition of the RL1 genes occurred in the common progenitor of all primate simplexviruses and was then lost during the divergence of the cercopithecine simplexviruses.

Gene prediction algorithms (GeneMark), trained with the codon usage of SA8 and B virus open reading frames, did not detect additional open reading frames in the SA8 genome. GeneMark analysis of the B virus genome had predicted a putative open reading frame UL53A, antisense to the UL53 open reading frame, that did not have an homologue in HSV-1 or HSV-2 (Perelygina et al., 2003). We investigated the corresponding region of the SA8 genome by performing a six frame translation and aligning the translated sequence with the predicted translated exons of the putative UL53A open reading frame of B virus. We found a considerable homology to the first exon that began with a possible start codon AUG. The homology extended to part of the second exon, but the intron/exon boundary was not clear by homology only. GeneMark analysis of this region also failed to support the existence of a potential UL53A gene in SA8. A considerable homology exists between SA8 and B virus at the DNA level in this region, which is also very G/C rich. It is therefore possible that the homology of the predicted protein sequences is simply the result of the DNA homology and not due to the presence of an open reading frame. Studies of the expression of the RNA in this region will resolve this issue.

GeneMark analysis failed to predict the open reading frame US11, which in fact can be easily predicted by homology with B virus, HSV-1, and HSV-2. However, the codon composition of US11 differs from all the other open reading frames of SA8, having a greater G/C content in the first (84.5%) and second position (77.6%) than in the third position (75.0%). All the other SA8 open reading frames, used to train GeneMark, have a higher G/C content in the third position (see below), and this may be the reason why US11 was not predicted. Interestingly, the same unique codon composition is found in the US11 open reading frame of B virus (data not shown).

The codon usage of the open reading frames (Table 2) maximizes the G/C content of the genome, as is the case for

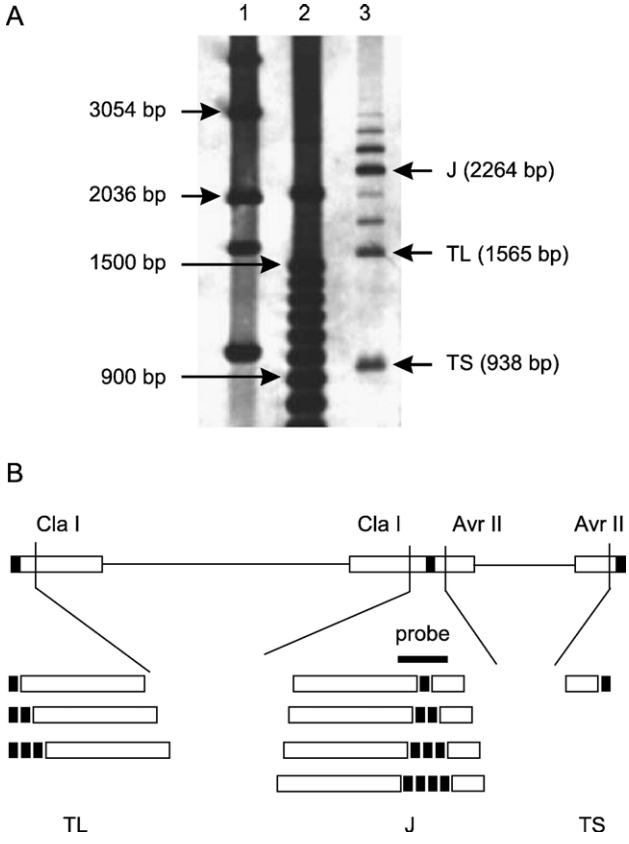

Fig. 4. Detection of multiple copies of the ‘a’ sequence. (A) Southern blot of *Avr*II/*Cla*I-digested DNA. Lanes 1 and 2 are 1 kb and 100 bp markers, respectively. Relevant fragment sizes have been indicated. Lane 3 is SA8 DNA digested with *Avr*II/*Cla*I and hybridized with a probe that spans the junction between long and short inverted repeats. The bands predicted for the terminal short (TS), terminal long (TL), and junction (J) fragments are indicated. (B) Map of the restriction fragments shown in panel A. The predicted sizes for the terminal long fragments are 1565, 1801, and 2037 bp with one, two, or three copies of the ‘a’ sequence, respectively; the predicted size for the terminal short fragment is 938 bp, and the predicted sizes for the junction fragment are 2264, 2500, 2736, and 2972 bp with one, two, three, or four copies of the ‘a’ sequence, respectively. The ‘a’ sequences (236 bp) are represented by the black boxes.

Table 1
ORF and other features of the SA8 genome

Feature name	Position		% G/C	Strand	Length (codons)	Protein homology (%)			Gene product predicted function[a]
	Start	End				B virus	HSV-1	HSV-2	
'a' sequence	1	236	88.5						Terminal direct repeat
TRL	237	9207	79.7						Terminal long repeat region
RL2			79.0	+	631	63.0	36.5	39.8	Immediate–early protein ICP0; multifunctional regulatory protein
Exon 1	2178	2225							
Exon 2	2448	3020							
Exon 3	3398	4669							
UL	9208	114322	74.5						Unique long region
UL1	9469	10167	66.2	+	233	65.2	48.7	46.9	Virion membrane glycoprotein L; in complex with gH; membrane fusion
UL2	10235	10990	72.4	+	252	86.1	72.9	71.7	Uracil–DNA glycosylase; DNA repair
UL3	11130	11813	74.5	+	228	74.9	52.9	52.4	Colocalization with ICP22 and UL4 in small, dense nuclear bodies
UL4	11897	12511	74.5	–	205	85.7	55.8	54.2	Colocalization with ICP22 and UL3 in small, dense nuclear bodies
UL5	12536	15163	69.3	–	876	93.3	80.9	81.6	Component of helicase–primase complex
UL6	15162	17198	75.6	+	679	90.3	65.4	67.0	Capsid protein; DNA cleavage/packaging
UL7	17149	18039	73.1	+	297	85.1	62.8	63.5	Capsid protein; DNA cleavage/packaging
UL8	18213	20489	80.7	–	759	77.6	55.7	58.5	Component of helicase–primase complex
UL9	23050	24402	72.2	+	451	90.7	75.7	78.0	Ori binding protein; helicase activity
UL10	23050	24402	74.6	+	451	86.7	64.2	63.6	Virion membrane glycoprotein M; proposed role in capsid envelopment
UL11	24649	24912	74.3	–	88	78.2	44.8	48.3	Myristylated tegument protein; capsid envelopment
UL12	24855	26702	76.8	–	616	79.8	62.1	60.8	DNase; endoexonuclease; processing of DNA replication intermediates
UL13	26702	28246	72.6	–	515	90.9	65.0	66.9	Virion protein kinase
UL14	28018	28662	73.4	–	215	90.7	64.0	68.2	Minor tegument protein
UL15			69.5	+	736	96.6	87.5	87.2	DNA cleavage/packaging; transiently associated with maturing capsids
Exon 1	28761	29789							
Exon 2	33312	34490							
UL16	29948	31036	77.0	–	363	87.8	62.7	64.6	Capsid associated; DNA cleavage/packaging; located in UL15 intron
UL17	31057	33168	79.3	–	704	86.7	65.9	67.7	Tegument protein; DNA cleavage/packaging
UL18	34642	35595	74.8	–	318	94.3	80.4	81.7	Capsid protein VP23; forms triplexes with VP19C that connect pentons and hexons in capsids
UL19	35743	39876	72.7	–	1378	95.5	86.8	86.5	Major capsid protein VP5; forms pentons and hexons of capsid shell
UL20	40057	40734	71.7	–	226	89.2	65.3	65.8	Virion membrane protein; virion egress; syn5 locus
UL21	41301	42881	74.9	+	527	85.9	61.2	66.0	Nucleotidylated phosphoprotein; interacts with microtubules and facilitates intracellular transport of the virus
UL22	43029	45620	77.0	–	864	81.6	57.4	59.7	Virion membrane glycoprotein H; in complex with gL; membrane fusion, entry, cell-to-cell spread
UL23	45797	46933	72.0	–	379	80.5	55.3	57.4	Thymidine kinase
UL24	46868	47650	71.2	+	261	83.8	63.1	61.9	Nonglycosylated membrane-associated protein; syn5 locus
UL25	47847	49580	75.0	+	578	91.3	79.7	80.1	Minor capsid protein; DNA packaging; possible role in DNA anchoring
UL26	49793	51577	78.7	+	595	76.4	60.4	58.2	Capsid maturation protease
UL26.5	50699	51577	80.2	+	293	66.1	43.8	49.7	Scaffolding protein
UL27	51943	54600	70.7	–	886	88.4	77.4	77.9	Virion membrane glycoprotein B; cell entry; contains syn3 locus
UL28	54631	56985	74.2	–	785	92.3	81.4	82.4	DNA cleavage/packaging; transiently associated with maturing capsids
UL29	57464	61054	71.5	–	1197	94.4	82.2	81.9	Single-strand DNA binding protein; key role in assembly of DNA replication proteins

Table 1 (*continued*)

Feature name	Position		% G/C	Strand	Length (codons)	Protein homology (%)			Gene product predicted function[a]
	Start	End				B virus	HSV-1	HSV-2	
oriL	61445	61542	47.8						Inverted repeat region of origin oriL
UL30	61764	65444	71.3	+	1227	90.1	77.2	77.7	DNA polymerase catalytic subunit; complexes with UL42
UL31	65389	66303	72.0	–	305	90.0	75.3	74.3	Nuclear phosphoprotein; interacts with UL34; capsid egress from nucleus
UL32	66296	68068	76.0	–	591	87.8	73.9	73.4	DNA packaging; not associated with capsids
UL33	68067	68474	70.6	+	136	88.6	71.5	71.5	DNA packaging; not associated with capsids
UL34	68524	69324	75.0	+	267	86.8	65.4	68.4	Type II nuclear membrane-associated phosphoprotein; interacts with UL31; capsid egress from nucleus
UL35	69442	69786	76.2	+	115	79.8	58.0	54.5	Basic phosphorylated capsid protein VP26
UL36	69965	79177	78.4	–	3071	79.9	58.2	58.7	Very large tegument protein; interacts with UL19 and UL37
UL37	79388	83041	76.9	–	1218	85.3	68.2	67.1	Minor tegument protein
UL38	83508	84887	76.8	+	460	89.7	68.8	66.9	Capsid protein VP19C, forms triplexes with VP23 that connect pentons and hexons in capsids
UL39	85252	88179	70.3	+	976	83.9	66.7	65.4	Large subunit of ribonucleotide reductase
UL40	88237	89214	66.4	+	326	89.5	83.4	83.4	Small subunit of ribonucleotide reductase
UL41	89390	90844	70.7	–	485	90.7	70.7	71.5	Tegument phosphoprotein; virion-associated host shutoff (vhs) protein
UL42	91272	92588	74.8	–	439	65.1	42.9	45.2	Double-stranded DNA binding protein, DNA polymerase subunit
UL43	92733	93878	80.0	+	382	72.3	42.8	42.5	Predicted membrane-associated protein
UL44	94112	95506	78.9	+	465	76.1	47.8	48.9	Virion membrane glycoprotein C; cell attachment; blocking host immune response
UL45	95717	96241	79.8	+	175	83.3	61.0	56.4	Type II membrane protein; possible role in cell fusion
UL46	96429	98471	76.8	–	681	83.7	53.2	54.7	Tegument phosphoprotein VP11/12; modulates alpha *trans*-inducing factor activity
UL47	98655	100679	79.3	–	675	81.8	58.3	54.2	Tegument phosphoprotein VP13/14; *O*-glycosylated; modulate–TIF activity; RNA binding
UL48	101176	102642	72.0	–	489	90.8	67.6	67.0	Major tegument protein VP16; *trans*-activator of genes
UL49	102951	103778	75.9	–	276	73.1	40.4	35.6	Major tegument protein VP22; binds RNA; carrier of mRNA from infected to uninfected cells
UL49.5	104165	104398	70.6	–	78	58.4	37.7	35.1	Envelope protein
UL50	104413	105516	75.9	+	368	83.1	55.9	53.1	Deoxyuridine triphosphatase
UL51	105673	106359	74.8	–	229	73.6	64.9	63.6	Capsid/tegument-associated phosphoprotein
UL52	106397	109558	75.2	+	1054	90.2	70.7	69.3	Component of helicase–primase complex
UL53	109510	110514	71.1	+	335	90.4	66.8	68.6	Membrane glycoprotein K; virion egress; contains syn1 locus
UL54	110790	112322	75.9	+	511	75.9	50.2	51.2	Immediate–early protein ICP27; regulates some early and all late gene expression
UL55	112612	113184	71.2	+	191	84.2	61.8	62.4	Nuclear matrix-associated protein
UL56	113344	114024	73.8	–	227	56.6	29.2	30.1	Type II membrane protein; involved in virus pathogenicity
IRL	114323	123293	79.7						Internal copy of large inverted repeat region
RL2			79.0		631	63.0	36.5	39.8	Immediate–early protein ICP0; multifunctional regulatory protein
Exon 3	118861	120132							
Exon 2	120510	121082							
Exon 1	121305	121352							
‘a’ sequence	123294	123529	88.5						Inverted copy of ‘a’ sequence
IRS	123530	129918	82.2						Internal copy of small inverted repeat region
RS1	125207	128707	84.1	–	1167	76.5	60.7	62.4	Immediate–early protein ICP4; regulator of gene expression
OriS	129452	129623	47.7						Inverted repeat region of replication origin oriS
US	129919	144385	76.3						Unique small region
US1	130088	131383	73.3	+	432	60.1	34.3	38.5	Immediate–early protein ICP22; required for optimal ICP0 expression

(continued on next page)

Table 1 (*continued*)

Feature name	Position		% G/C	Strand	Length (codons)	Protein homology (%)			Gene product predicted function[a]
	Start	End				B virus	HSV-1	HSV-2	
US2	131610	132512	78.8	–	301	78.0	54.3	50.2	Tegument protein
US3	132837	134201	73.8	+	455	82.6	59.3	60.1	Protein kinase; antiapoptotic activity
US4	134314	136128	80.7	+	605	58.6	17.6	33.1	Virion membrane glycoprotein G; entry into polarized cells
US5	136332	136652	78.2	+	107	54.7	27.2	17.4	Glycoprotein J; block apoptosis
US6	136982	138169	75.7	+	396	81.7	53.6	54.5	Virion membrane glycoprotein D; cell entry; interacts with cellular receptors
US7	138375	139574	77.6	+	400	68.2	43.8	46.5	Virion membrane glycoprotein I; in complex with gE; basolateral viral spread
US8	139886	141508	76.2	+	541	73.5	43.9	44.3	Virion membrane glycoprotein E; in complex with gI; basolateral viral spread
US8.5	141453	141761	76.7	+	103	79.4	41.2	45.1	Nucleolar phosphoprotein
US9	141852	142127	75.0	+	92	66.7	48.9	52.8	Tegument protein
US10	142577	143407	75.0	–	277	69.6	37.3	42.0	Tegument protein
US11	143205	143552	79.4	–	116	53.0	36.5	33.9	RNA binding tegument protein; interacts with protein kinase R
US12	143585	143821	70.5	–	79	57.7	29.5	25.6	Immediate–early protein ICP47; inhibits antigen presentation
TRS	144083	150479	82.2						Terminal copy of small repeat region
OriS	144386	144557	47.7						Inverted repeat region of replication origin oriS
RS1	145302	148805	84.1	+	1168	76.5	60.7	62.4	Immediate–early protein ICP4; regulator of gene expression
'a' sequence	150480	150715	88.5						Terminal direct repeat

[a] The predicted function of the SA8 gene products is the same as reported in Perelygina et al. (2003) for the homologous genes of B virus.

HSV (Hall et al., 1986; McGeoch et al., 1986; Schachtel et al., 1991) and for B virus (Perelygina et al., 2003). For example, the codon AUC coded for 93.2% of the occurrences of isoleucine, while the codons AUA and AUU, which do not contain C or G, accounted for only 6.8% of the occurrences. For serine, three codons, AGC, UCC, and UCG, each had the same content of G or C; together they accounted for 95.3% of the occurrences, with approximately the same frequency of 30.3%, 29.3%, and 35.7%, respectively. This rule had no exception in all 18 amino acids that are coded by more than one codon. The G/C frequency was 74.5%, 55.9%, and 91.4% for the first, second, and third codon positions, respectively.

Polyadenylation signals were sought by looking for the sequence AATAAA in both the sequence of SA8 and the published sequence of B virus (Perelygina et al., 2003). The positions of the polyadenylation signals are shown in Fig. 1. Putative transcriptional units of SA8 were identical to those of the B virus. SA8 and HSV polyadenylation signals associated with open reading frames showed only three differences: (i) there was no polyadenylation signal after UL24 in SA8; (ii) in SA8 there was no polyadenylation signal downstream of UL53, suggesting that this gene is transcribed with UL52 and UL54; and (iii) in SA8 there was a polyadenylation signal after US5, suggesting that US3, US4, and US5 are transcribed as one unit. In HSV this polyadenylation signal was downstream of US4, suggesting that US5 is transcribed along with US6 and US7 (see Fig. 1). In addition, the SA8 genome contained several polyadenylation signals that are apparently not associated with any known open reading frame. Two were in the LAT region of the long repeat and three were closely spaced in the short repeat (as shown in Fig. 1). One signal was on the untranscribed strand complementary to the UL52 open reading frame.

We have also scanned the two strands of the SA8 genome for the presence of nine variants of the polyadenylation signal, as described by Beaudoing et al. (2000). Variant polyadenylation signals were found near the 3′ end of open reading frames UL9 and UL20 (TATAAA), UL19, US12 and US8/8.5 (CATAAA), and UL34 (GATAAA). However, the functional significance of these putative polyadenylation signals is doubtful since 69 more sequences matching variant signals were found scattered in the SA8 genome, with no obvious association with any open reading frame.

The latency-associated transcript (LAT) region is located in the inverted repeat region of simplexviruses, 3′ of the RL2 gene. Studies on HSV-1 have shown that this region produces a family of RNA molecules antisense to the RL2 gene, which are the only transcripts found in latently infected neurons. LATs have been implicated in the establishment of latency and reactivation, survival of infected neuronal cells, and inhibition of apoptosis (for recent reviews, see Jones, 2003; Wagner and Bloom, 1997). In HSV-1, the major transcript is approximately 8.5 kb long, running antisense to RL2 in the long inverted repeat, across the L/S junction, and into the short inverted

repeat region, where its 3′ terminus overlaps with the RS1 (Devi-Rao et al., 1991; Dobson et al., 1989). We have aligned the sequence of the 8.5-kb LAT of HSV-1 with the SA8 inverted repeat region and found considerable homology in the region of the third and second exon of RL2, reflecting homology of the amino acid sequence. DNA homology was less obvious upstream and downstream of RL2. In particular, it was not possible to find significant homologies between the LAT promoter region of HSV-1 (Chen et al., 1995; Dobson et al., 1989; Goins et al., 1994) and the corresponding region of the SA8 genome. At the 3′ end of the LAT region on the short inverted repeat, three polyadenylation signals were detected in the homologous position of the LAT polyadenylation signal of HSV-1. The positions of the putative 8.5-kb LAT homologue and the polyadenylation signals are shown on Fig. 1.

Conclusions

The complete sequencing of the SA8 genome has revealed a close homology with the monkey B virus. HSV-1 and HSV-2 also share a close homology with SA8, but it is obvious that a common progenitor of HSV-1 and HSV-2 diverged from a common progenitor of the monkey B virus and SA8. Particularly interesting is the fact that both B virus and SA8 lack the RL1 open reading frame, or neurovirulence gene, found in HSV-1 and HSV-2, suggesting that human simplexviruses may have a different pathogenetic mechanism than monkey simplexviruses. The availability of a complete sequence for the four simplexviruses will facilitate the study of these pathogenetic mechanisms.

Materials and methods

Viruses and cells

SA8 virus strain B264 (ATCC number VR-936) was cultured in Vero cell line (ATCC number CCL-81). SA8 DNA was prepared from confluent Vero cells that were infected at a multiplicity of infection of 1 and incubated until cytopathic effect was close to 100%. Cells were lysed in 10 mM EDTA, 10 mM Tris pH 8.0, 1% SDS, and the lysate was incubated with 200 μg of proteinase K (Sigma) at 37 °C for 24 h. The lysate was then extracted three times with phenol/chloroform and the DNA was precipitated with 100 mM NaCl and two volumes of ethanol. The DNA pellet was resuspended in TE buffer (1 mM EDTA, 10 mM Tris, pH 8) and separated by ultracentrifugation on a NaI gradient, with 0.5 μg/ml of ethidium bromide, as previously described (Walboomers and Schegget, 1976). Because of the high G/C content, SA8 DNA formed a distinct band that sedimented at a higher density in the NaI gradient compared

Table 2
Codon usage in SA8 open reading frames

Amino acid	Codon	*n*	%
Ala	GCA	65	1.1
	GCC	3219	53.5
	GCG	2652	44.0
	GCU	86	1.4
Arg	AGA	37	1.0
	AGG	140	3.8
	CGA	134	3.6
	CGC	2040	55.4
	CGG	1262	34.3
	CGU	67	1.8
Asn	AAC	610	96.5
	AAU	22	3.5
Asp	GAC	1928	95.8
	GAU	85	4.2
Cys	UGC	596	91.3
	UGU	57	8.7
Gln	CAA	32	3.8
	CAG	800	96.2
Glu	GAA	105	5.2
	GAG	1907	94.8
Gly	GGA	133	4.5
	GGC	1589	53.3
	GGG	1181	39.6
	GGU	79	2.6
His	CAC	819	97.5
	CAU	21	2.5
Ile	AUA	20	2.6
	AUC	710	93.2
	AUU	32	4.2
Leu	CUA	27	0.7
	CUC	1359	37.2
	CUG	2172	59.4
	CUU	42	1.1
	UUA	7	0.2
	UUG	49	1.3
Lys	AAA	34	7.9
	AAG	397	92.1
Phe	UUC	1022	83.5
	UUU	202	16.5
Pro	CCA	59	1.7
	CCC	1946	56.3
	CCG	1375	39.8
	CCU	76	2.2
Ser	AGC	570	30.3
	AGU	23	1.2
	UCA	12	0.6
	UCC	550	29.3
	UCG	670	35.7
	UCU	54	2.9
Stop	UAA	12	16.0
	UAG	24	32.0
	UGA	39	52.0
Thr	ACA	22	1.3
	ACC	863	49.4
	ACG	839	48.0
	ACU	23	1.3
Tyr	UAC	828	94.2
	UAU	51	5.8
Val	GUA	19	0.7
	GUC	1224	48.2
	GUG	1225	48.2

to cellular DNA. The SA8 DNA band was collected through a wide bore needle, extracted three times with butanol to remove ethidium bromide, and dialyzed overnight against TE buffer with 100 mM NaCl. The DNA was then ethanol precipitated and used for sequencing.

Sequencing strategy

A random shotgun library was prepared from the SA8 genomic DNA using the TOPO Shotgun Subcloning Kit (Invitrogen). Briefly, the DNA was randomly sheared using the provided nebulizer and fragments in the 1- to 3-kb range were isolated by agarose gel electrophoresis. These fragments were blunt ended by incubating in the presence of Klenow DNA polymerase and T4 DNA polymerase, dephosphorylated with calf intestinal phosphatase, and ligated into the vector pCR4 Blunt-TOPO (Invitrogen). The ligation products were electroporated into TOP10 electrocompetent *Escherichia coli* (Invitrogen), β galactosidase-negative recombinants (white colonies on X-gal plates) were selected and then grown in 96-well blocks. Plasmid DNA was prepared using the Wizard SV96 kit (Promega). All sequencing was performed on ABI3100 Sequencers (ABI) using BigDye 3.1 chemistry or a 3:1 mix of BigDye 3.1 and dGTP chemistry with the addition of enhancer A (Invitrogen). In order to fill gaps in sequencing left after the shotgun approach, numerous custom primers were employed to extend sequence data from specific clones or to generate PCR products for sequencing. In some instances, specific regions were subcloned into pBlueScript II KS (Stratagene) to isolate "problem" regions and facilitate sequencing through these regions. During the assembly process, the sequence data were periodically screened against GenBank to remove contaminating sequences that were of human origin.

In order to confirm the boundary of the repeat regions, clones that had been identified to span US1 or US12 and the short repeat unit and UL1 or UL56 and the long repeat unit were completely sequenced. By combining this data with the terminal sequences, the final genome arrangement was inferred. The sequencing data that were assembled into the internal copy of the repeats were closely scrutinized to determine if discrepancies were present, which would indicate that the inverted repeats were not "perfect" repeats; however, none were found. The sequence redundancy in this region averaged 10- to 15-fold coverage, with a 6.4-fold redundancy for the genome overall.

In order to identify the genomic termini, genomic DNA was end repaired with T4 DNA polymerase and Klenow DNA polymerase prior to digestion with *Cla*I, for the long terminal sequence, or *Avr*II for the short terminal sequence. The digested DNA was cloned into pBluescript (Stratagene) prepared with *Cla*I/*Sma*I for the *Cla*I-digested DNA or *Xba*I/*Eco*RV for the *Avr*II-digested DNA. All enzymes used for cloning the termini were obtained from New England Biolabs. The resulting clones were mapped to identify the ones that contained inserts consistent with the predicted size of the terminal fragments, and at least six independent clones were sequenced for each end.

Southern blot

Genomic DNA was double digested with the restriction enzymes *Avr*II and *Cla*I (New England Biolabs). The fragments were separated on a 1.5% agarose gel and transferred to nylon membrane (Hybond N+, Amersham Biosciences). The membrane was probed with a PCR product spanning the internal junction and developed using the ECL Direct Labeling and Detection System (Amersham Biosciences).

Sequence analysis

Sequence data were compiled using the Staden Package (Staden et al., 1999) and subsequent analysis of the sequence was conducted using various software packages that included DNAStar (LaserGene), Vector NTI (InforMax), Artemis (Rutherford et al., 2000), as well as NCBI and other Web-based tools. Coding region predictions were performed using the programs GeneMark 2.4 and MkMat 2.0 (GeneProbe), which are based on the GeneMark/hmm algorithm (Besemer and Borodovsky, 1999; Besemer et al., 2001; Lukashin and Borodovsky, 1998). Tandem repeat sets were identified using the program Tandem Repeats Finder (Benson, 1999).

Acknowledgments

This work was supported in part by a grant from the Health Canada Genomic Research and Development Program. The support of the DNA Core Facility, National Microbiology Laboratory was invaluable for the completion of this project.

References

Barnett, B.C., Dolan, A., Telford, E.A., Davison, A.J., 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.

Beaudoing, E., Freier, S., Wyatt, J.R., Claverie, J.M., Gautheret, D., 2000. Patterns of variant polyadenylation signal usage in human genes. Genome Res. 10, 1001–1010.

Benson, G., 1999. Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. 27, 573–580.

Besemer, J., Borodovsky, M., 1999. Heuristic approach to deriving models for gene finding. Nucleic Acids Res. 27, 3911–3920.

Besemer, J., Lomsadze, A., Borodovsky, M., 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. 29, 2607–2618.

Borchers, K., Ozel, M., 1993. Simian agent 8 (SA8): morphogenesis and ultrastructure. Zentralbl. Bakteriol. 279, 526–536.

Borchers, K., Ozel, M., Pauli, G., Gelderblom, H.R., Ludwig, H., 1990. Conserved epitopes of simian herpesvirus SA 8 and bovine herpesvirus type 2. Arch. Virol. 111, 1–14.

Brack, M., Eichberg, J.W., Heberling, R.L., Kalter, S.S., 1985. Experimental Herpes neonatalis in SA 8-infected baboons (*Papio cynocephalus*). Lab. Anim. 19, 125–131.

Cassady, K.A., Gross, M., Roizman, B., 1998. The herpes simplex virus US11 protein effectively compensates for the gamma1(34.5) gene if present before activation of protein kinase R by precluding its phosphorylation and that of the alpha subunit of eukaryotic translation initiation factor 2. J. Virol. 72, 8620–8626.

Chen, X., Schmidt, M.C., Goins, W.F., Glorioso, J.C., 1995. Two herpes simplex virus type 1 latency-active promoters differ in their contributions to latency-associated transcript expression during lytic and latent infections. J. Virol. 69, 7899–7908.

Chou, J., Roizman, B., 1994. Herpes simplex virus 1 gamma(1)34.5 gene function, which blocks the host response to infection, maps in the homologous domain of the genes expressed during growth arrest and DNA damage. Proc. Natl. Acad. Sci. U.S.A. 91, 5247–5251.

Chou, J., Kern, E.R., Whitley, R.J., Roizman, B., 1990. Mapping of herpes simplex virus-1 neurovirulence to (134.5, a gene nonessential for growth in culture. Science 250, 1262–1266.

Chou, J., Chen, J.J., Gross, M., Roizman, B., 1995. Association of a Mr 90,000 phosphoprotein with protein kinase PKR in cells exhibiting enhanced phosphorylation of translation initiation factor eIF-2 alpha and premature shutoff of protein synthesis after infection with gamma 134.5- mutants of herpes simplex virus 1. Proc. Natl. Acad. Sci. U.S.A. 92, 10516–10520.

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

Davidson, W.L., Hummeler, K., 1960. B virus infection in man. Ann. N. Y. Acad. Sci. 85, 970–979.

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

Devi-Rao, G.B., Goodart, S.A., Hecht, L.M., Rochford, R., Rice, M.K., Wagner, E.K., 1991. Relationship between polyadenylated and non-polyadenylated herpes simplex virus type 1 latency-associated transcripts. J. Virol. 65, 2179–2190.

Dobson, A.T., Sederati, F., Devi-Rao, G., Flanagan, W.M., Farrell, M.J., Stevens, J.G., Wagner, E.K., Feldman, L.T., 1989. Identification of the latency-associated transcript promoter by expression of rabbit beta-globin mRNA in mouse sensory nerve ganglia latently infected with a recombinant herpes simplex virus. J. Virol. 63, 3844–3851.

Dolan, A., McKie, E., MacLean, A.R., McGeoch, D.J., 1992. Status of the ICP34.5 gene in herpes simplex virus type 1 strain 17. J. Gen. Virol. 73, 971–973.

Dolan, A., Jamieson, F.E., Cunningham, C., Barnett, B.C., McGeoch, D.J., 1998. The genome sequence of herpes simplex virus type 2. J. Virol. 72, 2010–2021.

Eberle, R., Black, D., 1991. The simian herpesvirus SA8 homologue of the herpes simplex virus gB gene: mapping, sequencing, and comparison to the HSV gB. Arch. Virol. 118, 67–86.

Eberle, R., Black, D., 1993. Sequence analysis of herpes simplex virus gB gene homologs of two platyrrhine monkey alpha-herpesviruses. Arch. Virol. 129, 167–182.

Eberle, R., Black, D., Hilliard, J.K., 1989. Relatedness of glycoproteins expressed on the surface of simian herpes-virus virions and infected cells to specific HSV glycoproteins. Arch. Virol. 109, 233–252.

Eberle, R., Zhang, M., Black, D.H., 1993. Gene mapping and sequence analysis of the unique short region of the simian herpesvirus SA 8 genome. Arch. Virol. 130, 391–411.

Eberle, R., Black, D.H., Lipper, S., Hilliard, J.K., 1995. Herpesvirus papio 2, an SA8-like alpha-herpesvirus of baboons. Arch. Virol. 140, 529–545.

Eichberg, J.W., McCullough, B., Kalter, S.S., Thor, D.E., Rodriguez, A.R., 1976. Clinical, virological, and pathological features of herpesvirus SA8 infection in conventional and gnotobiotic infant baboons (*Papio cynocephalus*). Arch. Virol. 50, 255–270.

Everett, R.D., Fenwick, M.L., 1990. Comparative DNA sequence analysis of the host shutoff genes of different strains of herpes simplex virus: type 2 strain HG52 encodes a truncated UL41 product. J. Gen. Virol. 71, 1387–1390.

Fornace Jr., A.J., Nebert, D.W., Hollander, M.C., Luethy, J.D., Papathanasiou, M., Fargnoli, J., Holbrook, N.J., 1989. Mammalian genes coordinately regulated by growth arrest signals and DNA-damaging agents. Mol. Cell. Biol. 9, 4196–4203.

Goins, W.F., Sternberg, L.R., Croen, K.D., Krause, P.R., Hendricks, R.L., Fink, D.J., Straus, S.E., Levine, M., Glorioso, J.C., 1994. A novel latency-active promoter is contained within the herpes simplex virus type 1 UL flanking repeats. J. Virol. 68, 2239–2252.

Hall, J.D., Gibbs, J.S., Coen, D.M., Mount, D.W., 1986. Structural organization and unusual codon usage in the DNA polymerase gene from herpes simplex virus type 1. DNA 5, 281–288.

Hazuda, D.J., Perry, H.C., Naylor, A.M., McClements, W.L., 1991. Characterization of the herpes simplex virus origin binding protein interaction with oriS. J. Biol. Chem. 266, 24621–24626.

Jackson, S.A., DeLuca, N.A., 2003. Relationship of herpes simplex virus genome configuration to productive and persistent infections. Proc. Natl. Acad. Sci. U.S.A. 100, 7871–7876.

Jones, C., 2003. Herpes simplex virus type 1 and bovine herpesvirus 1 latency. Clin. Microbiol. Rev. 16, 79–95.

Kalter, S.S., Weiss, S.A., Heberling, R.L., Guajardo, J.E., Smith III, G.C., 1978. The isolation of herpesvirus from trigeminal ganglia of normal baboons (*Papio cynocephalus*). Lab. Anim. Sci. 28, 705–709.

Katz, D., Shi, W., Krug, P.W., Henkel, R., McClure, H., Hilliard, J.K., 2002. Antibody cross-reactivity of alphaherpesviruses as mirrored in naturally infected primates. Arch. Virol. 147, 929–941.

Levin, J.L., Hilliard, J.K., Lipper, S.L., Butler, T.M., Goodwin, W.J., 1988. A naturally occurring epizootic of simian agent 8 in the baboon. Lab. Anim. Sci. 38, 394–397.

Lord, K.A., Hoffman-Liebermann, B., Liebermann, D.A., 1990. Sequence of MyD116 cDNA: a novel myeloid differentiation primary response gene induced by IL6. Nucleic Acids Res. 18, 2823.

Lukashin, A.V., Borodovsky, M., 1998. GeneMark.hmm: new solutions for gene finding. Nucleic Acids Res. 26, 1107–1115.

Malherbe, H., Harwin, R., 1958. Neurotropic virus in African Monkeys. Lancet 2, 530.

Malherbe, H., Harwin, R., 1963. The cytopathic effects of vervet monkey viruses. S. Afr. Med. J. 37, 407–411.

Malherbe, H., Strickland-Cholmley, M., 1969a. Simian herpesvirus SA8 from a baboon. Lancet 2, 1427.

Malherbe, H., Strickland-Cholmley, M., 1969b. Virus from baboons. Lancet 2, 1300.

Martino, M.A., Hubbard, G.B., Butler, T.M., Hilliard, J.K., 1998. Clinical disease associated with simian agent 8 infection in the baboon. Lab. Anim. Sci. 48, 18–22.

McGeoch, D.J., Barnett, B.C., 1991. Neurovirulence factor. Nature 353, 609.

McGeoch, D.J., Dolan, A., Donald, S., Rixon, F.J., 1985. Sequence determination and genetic content of the short unique region in the genome of herpes simplex virus type 1. J. Mol. Biol. 181, 1–13.

McGeoch, D.J., Dolan, A., Donald, S., Brauer, D.H., 1986. Complete DNA sequence of the short repeat region in the genome of herpes simplex virus type 1. Nucleic Acids Res. 14, 1727–1745.

McGeoch, D.J., Moss, H.W., McNab, D., Frame, M.C., 1987. DNA sequence and genetic content of the *Hin*dIII l region in the short unique component of the herpes simplex virus type 2 genome: identification of the gene encoding glycoprotein G, and evolutionary comparisons. J. Gen. Virol. 68, 19–38.

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

McGeoch, D.J., Cunningham, C., McIntyre, G., 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.

McVoy, M.A., Nixon, D.E., Adler, S.P., Mocarski, E.S., 1998. Sequences within the herpesvirus-conserved pac1 and pac2 motifs are required for cleavage and packaging of the murine cytomegalovirus genome. J. Virol. 72, 48–56.

McVoy, M.A., Nixon, D.E., Hur, J.K., Adler, S.P., 2000. The ends on herpesvirus DNA replicative concatemers contain pac2 cis cleavage/packaging elements and their formation is controlled by terminal cis sequences. J. Virol. 74, 1587–1592.

Mocarski, E.S., 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.

Mohr, I., Gluzman, Y., 1996. A herpesvirus genetic element which affects translation in the absence of the viral GADD34 function. EMBO J. 15, 4759–4766.

Mulvey, M., Poppers, J., Ladd, A., Mohr, I., 1999. A herpesvirus ribosome-associated, RNA-binding protein confers a growth advantage upon mutants deficient in a GADD34-related function. J. Virol. 73, 3375–3385.

Mulvey, M., Camarena, V., Mohr, I., 2004. Full resistance of herpes simplex virus type 1-infected primary human cells to alpha interferon requires both the Us11 and gamma(1)34.5 gene products. J. Virol. 78, 10193–10196.

Perelygina, L., Zhu, L., Zurkuhlen, H., Mills, R., Borodovsky, M., Hilliard, J.K., 2003. Complete sequence and comparative analysis of the genome of herpes B virus (cercopithecine herpesvirus 1) from a rhesus monkey. J. Virol. 77, 6167–6177.

Perry, L.J., McGeoch, D.J., 1988. The DNA sequences of the long repeat region and adjoining parts of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. 69, 2831–2846.

Roizman, B., 1979. The structure and isomerization of herpes simplex virus genome. Cell 16, 481–494.

Roizman, B., Pellet, P.E., 2001. The family *Herpesviridae*: a brief introduction. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott, Williams and Wilkins, pp. 2381–2397.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., Barrell, B., 2000. Artemis: sequence visualization and annotation. Bioinformatics 16, 944–945.

Sabin, A.B., 1949. Fatal B virus encephalomyelitis in a physician working with monkeys. J. Clin. Invest. 28, 808.

Sabin, A.B., Wright, A.M., 1934. Acute ascending myelitis following a monkey bite, with the isolation of a virus capable of reproducing the disease. J. Exp. Med. 59, 115–136.

Schachtel, G.A., Bucher, P., Mocarski, E.S., Blaisdell, B.E., Karlin, S., 1991. Evidence for selective evolution in codon usage in conserved amino acid segments of human alphaherpesvirus proteins. J. Mol. Evol. 33, 483–494.

Smiley, J.R., Duncan, J., Howes, M., 1990. Sequence requirements for DNA rearrangements induced by the terminal repeat of herpes simplex virus type 1 KOS DNA. J. Virol. 64, 5036–5050.

Staden, R., Beal, K.F., Bonfield, J.K., 1999. The Staden Package, 1998, vol. 132. The Humana Press, Inc., Totowa, NJ, pp. 115–130.

van Regenmortel, M.H.V., Fauquet, C.M., Bishop, D.H.L., Carstens, E.B., Estes, M.K., Lemon, S.M., Maniloff, J., Mayo, M.A., McGeoch, D.J., Pringle, C.R., Wickner, R.B., 2000. Virus taxonomy. VIIth Report of the ICTV. Academic Press, San Diego.

Varmuza, S.L., Smiley, J.R., 1985. Signals for site-specific cleavage of HSV DNA: maturation involves two separate cleavage events at sites distal to the recognition sequences. Cell 41, 793–802.

Veit, M., Ponimaskin, E., Baiborodin, S., Gelderblom, H.R., Schmidt, M.F., 1996. Intracellular compartmentalization of the glycoprotein B of herpesvirus Simian agent 8 expressed with a baculovirus vector in insect cells. Arch. Virol. 141, 2009–2017.

Wagner, E.K., Bloom, D.C., 1997. Experimental investigation of herpes simplex virus latency. Clin. Microbiol. Rev. 10, 419–443.

Walboomers, J.M., Schegget, J.T., 1976. A new method for the isolation of herpes simplex virus type 2 DNA. Virology 74, 256–258.

Weigler, B.J., 1992. Biology of B virus in macaque and human hosts: a review. Clin. Infect. Dis. 14, 555–567.

Whitley, R.J., Kern, E.R., Chatterjee, S., Chou, J., Roizman, B., 1993. Replication, establishment of latency, and induced reactivation of herpes simplex virus $\gamma_1$34.5 deletion mutants in rodent models. J. Clin. Invest. 91, 2837–2843.

Yao, X.D., Matecic, M., Elias, P., 1997. Direct repeats of the herpes simplex virus a sequence promote nonconservative homologous recombination that is not dependent on XPF/ERCC4. J. Virol. 71, 6842–6849.

Zhan, Q., Lord, K.A., Alamo Jr., I., Hollander, M.C., Carrier, F., Ron, D., Kohn, K.W., Hoffman, B., Liebermann, D.A., Fornace Jr., A.J., 1994. The gadd and MyD genes define a novel set of mammalian genes encoding acidic proteins that synergistically suppress cell growth. Mol. Cell. Biol. 14, 2361–2371.

Zhu, L., Perelygina, L., Patrusheva, I., Hilliard, J.K., 2004. Counteraction of PKR by B virus US11 protein. 23rd Annual Meeting of the American Society for Virology, W51-2 (Abstract).