The simian varicella virus genome contains an invertible 665 base pair terminal element that is absent in the varicella zoster virus genome

Ravi Mahalingam a,⁎, Wayne L. Gray b

a Department of Neurology, Mail Box B183, University of Colorado Health Sciences Center, 4200 East Ninth Avenue, Denver, CO 80262, USA
b Department of Microbiology and Immunology, University of Arkansas for Medical Sciences, Little Rock, AR 72205, USA

Received 19 March 2007; returned to author for revision 21 April 2007; accepted 26 April 2007
Available online 6 June 2007

Abstract

Simian varicella virus (SVV) causes chickenpox in monkeys, establishes latency and reactivates to produce zoster thus providing a model to study human varicella zoster virus (VZV) infection. Sequence analysis of a recombinant cosmid clone containing the left end of the SVV genome revealed a 665 base pair (bp) segment that is absent in VZV DNA. This segment inverts and contains 507 bp of unique sequences flanked on either side by 79 bp inverted repeats, making the SVV genome to be 124,785 bp in size. Part of the inverted repeat sequence (64 bp) is also present at the junction of the long and short segments of the SVV genome. The terminal DNA sequences are conserved among different SVV isolates and present in tissues from infected monkeys. The terminal region is transcriptionally active and is also present in the genomes of other animal varicelloviruses but absent in the VZV genome.

Keywords: Simian varicella virus; Varicella zoster virus; DNA sequence

Introduction

Simian varicella virus (SVV) is an alphaherpesvirus that causes varicella in non-human primates, establishes latent infection in ganglionic neurons (Mahalingam et al., 2002; Kennedy et al., 2004) and reactivates to produce zoster. Although varicella zoster virus (VZV) reactivation in humans (zoster) is generally localized to 1–3 dermatomes, SVV reactivation often appears as a whole-body rash. SVV shares virological, immunological, and pathological features with VZV and has provided a useful model to study varicella pathogenesis, latency and reactivation (Gray, 2004). The two virus genomes share a high degree of nucleotide homology (Gray and Oakes, 1984). Earlier, as part of our efforts to determine the complete nucleotide sequence of the SVV genome (Gray et al., 2001), we used oligonucleotide primers located at the leftward and the rightward ends of the virus genome to amplify the DNA segment (containing the leftward end of the virus genome) from concatemeric SVV genomes. We showed that SVV and VZV genomes differed at the leftward terminus: SVV lacks a VZV ORF 2 homologue and encodes an 879 base pair (bp) ORF A that is absent in VZV but has homology to VZV ORF 4. We also detected an 8-bp inverted repeat sequence flanking the unique long segment of the SVV genome (Mahalingam et al., 2000). To confirm these results independently, herein, we determined the sequence of the ends of the insert of a recombinant cosmid clone that contained the leftward end of SVV genome. Co-transfection of permissive cells with this recombinant cosmid clone along with three other overlapping recombinant cosmid clones that span the rest of the SVV genome produces infectious SVV (Gray and Mahalingam, 2005). We present our analysis of the sequence of the leftward end of SVV genome located within the recombinant cosmid clone and provide confirmation that these sequences are present in the virus genome.

Results

Identification of inverted repeat sequences at the leftward end of the SVV genome

We prepared cosmid clones spanning the SVV genome as described (Gray and Mahalingam, 2005). DNA sequence analysis
of SVV recombinant cosmid A, which includes 32.7 kilobases (kb) of SVV DNA located at the leftward end of the viral genome, using primer pWEBP1 (see Materials and methods section), revealed the nucleotides located at the left terminus of the SVV genome. Additional sequence analysis of cosmid A using SVV-specific primers indicated that the SVV left end includes a 665-bp terminal element composed of 79 bp inverted repeat sequences (TRL and IRL-A) flanking 507 bp of unique sequences (Figs. 1 and 2). SVV sequences that were located outside of the inverted repeats (nucleotides 666–761 in Fig. 2) overlapped with the leftward end of the SVV genome that we had previously published (Gray et al., 2001). Comparison of the sequences of the inverted repeats with the rest of the previously published SVV genomic sequences showed that of the 79 bp, 64 nucleotides were also located at the junction of the unique long (UL) and inverted repeat (IRS) of the SVV genome (IRL-B in Fig. 1). These results indicate that the 104,104 bp SVV UL component is bracketed by 64 bp inverted repeat sequences.

To confirm that these additional SVV sequences obtained using the recombinant cosmid clone were part of the SVV genome, we amplified, using PCR, a 404-bp DNA fragment that spanned the junction between the rightward and the leftward

---

**Fig. 1. Structure of SVV genome.** The 124.7-kb SVV genome consists of a 104.1-kb unique long (UL) and a 4.9-kb unique short (US) segment. The US segment is flanked by 7.5-kb terminal (TRS) and internal (IRS) repeat sequences. The location of the cosmid DNA insert (cosmid A) used in the identification of left end of the SVV genome is shown. Sequence analysis of the leftward end of cosmid A revealed the presence of a 507-bp unique sequence flanked by 79 bp inverted repeats (TRL and IRL-A). Of the 79 bp of the inverted repeat DNA sequences, 64 bp are also present at the other end of the UL segment (IRL-B). The remaining 15 bp of the repeat sequences are present only at the leftward end.

**Fig. 2. Nucleotide sequence at the left end of the SVV genome.** Sequences are shown starting from the left most end of the SVV genome (nucleotide 1). Location and the direction of the terminal (TRL) and internal (IRL-A) repeat regions (dark arrows), unique long segment (UL) and the primers (P1–P9) (thin arrows) used for the analysis are shown. The dashed arrows indicate the 15-bp portion of the repeat segment that is present only at the leftward end. The shaded region (46 nucleotides) is present as a direct repeat in MLMV and LVV strains of SVV (Fig. 5).
ends of the SVV genome, using primers P10 and P11 (Figs. 2 and 3) and DNA extracted from SVV-infected Vero cells in culture. Sequence analysis of the PCR fragment identified the leftward terminus of the SVV genome (Fig. 3), which was in agreement with the sequence information obtained using the cosmid clone (Fig. 2).

**Inversion of the terminal element located at the leftward end of the SVV genome**

To determine if the terminal element containing the inverted repeat sequences located at the leftward end of the SVV genome undergoes inversion in virus-infected cells in culture, we used primer P6 (located in the UL segment) in combination with either primer P1 or P4 on DNA extracted from SVV-infected Vero cells in PCR and analyzed the products by gel electrophoresis (Fig. 4). In one orientation (A in Fig. 4), primers P1 and P6 will generate a 654-bp PCR fragment and in the other orientation (B in Fig. 4), primers P4 and P6 will generate a 523-bp PCR fragment. The detection of the expected PCR fragments by agarose gel electrophoresis confirmed that the terminal element located at the leftward end of the SVV genome inverts relative to the SVV UL component (Fig. 4).
The terminal element located at the leftward end of the SVV genome is conserved among SVV strains, but varies in size

Our results demonstrated the presence of the terminal element in the genome of the SVV prototype virus, Delta herpesvirus (DHV). To determine if the terminal element is conserved between different SVV strains, we extracted total DNA from Vero cells infected with DHV, Hazelton herpesvirus (HAZV), Liverpool vervet virus (LVV) and Medical Lake Macaque virus (MLMV) (Gray and Gusick, 1996) and amplified the DNA sequences located at the leftward end of the virus genomes by PCR using primers P1 and P6 (Fig. 2). The PCR fragments derived from the DHV and the HAZV strains co-migrated on agarose gel, at around 650 bp, while the fragments derived from LVV and the MLMV co-migrated at around 700 bp (Fig. 5A). These results indicated that while the terminal element was present in all four of the SVV strains, they differed in their sizes. We cloned the amplified DNA fragments and determined the nucleotide sequence. The sequence of the region of the leftward ends that differed between SVV strains is presented in Fig. 5B. The DNA sequence of the terminal element of the DHV and the HAZV strains were identical, whereas in the LVV and the MLMV strains a 46-bp segment of unique sequences was present in two copies (Figs. 2 and 5B).

Detection of the terminal element in DNA from tissues from SVV-infected monkeys

To determine if the terminal element can be detected in tissues from monkeys that were intratracheally inoculated as well as naturally infected with SVV, we analyzed DNA extracted from blood mononuclear cells 14 days p.i. (acute, intratracheal), from lung 57 days p.i. (intratracheal), from lumbar ganglia 85 days after natural exposure and from trigeminal ganglia 57 days p.i. (intratracheal) by nested-PCR amplification using primers P1 and P5 (primary) and P2 and P4 (nest) followed by Southern blot hybridization using radioactively labeled oligonucleotide probe P3 (Fig. 2). We detected the expected 279-bp nested PCR amplification product in DNA extracted from all the tissues derived from SVV-infected monkeys (Fig. 6). These results further confirm that the terminal element is part of the infectious SVV genome.

Comparison of DNA sequences at the leftward end of different varicelloviruses

Although VZV genome contains an 88-bp sequence at the leftward end that is present as an inverted repeat at the junction of the long and short segments, it does not contain a terminal element similar to that seen in SVV (Davison and Scott, 1986). Equine herpesvirus (EHV) types 1 and 4 and pseudorabies virus (PRV) genomes have been sequenced and shown to contain terminal elements similar to that of SVV at their leftward ends with 87, 86 and 82 bp inverted repeats, respectively (Klupp et al., 2004; Telford et al., 1992, 1998). Comparative analysis of the inverted repeat sequences at the leftward end of SVV genome with those of other varicelloviruses, including VZV,
revealed the presence of conserved sequences (Fig. 7). EHV-1, EHV-4 and SVV genomes contained GGCC at the leftward terminus while VZV and PRV genomes contained A and GA preceding GGCC, respectively. In addition, a stretch of GC-rich sequences followed by a GAAAAAAA sequence and a putative CG motif (except in EHV-4) were also conserved within the first 90 nucleotides. Comparative analysis of the unique sequences within the terminal element did not reveal any conserved sequences (data not shown).

**Transcriptional activity at the leftward end of the SVV genome**

To determine if the terminal element codes for viral transcripts, RT-PCR was conducted using total RNA from SVV-infected Vero cells and primer pairs P1 and P7 or P8 and P9 (Fig. 2). We did not detect any RT-PCR product using primers P1 and P7. However, a 210-bp cDNA product was amplified using primers P8 and P9 (Fig. 8). The sequence of the 210-bp RT-PCR product was found to be identical to the portion of the SVV DNA sequence presented in Fig. 2. A subsequent study will characterize the transcript in detail.

Control reactions conducted without reverse transcriptase confirmed that the amplified product was not due to the presence of contaminating DNA. These results indicated that the terminal element located at the leftward end of SVV is transcriptionally active.

**Discussion**

This study demonstrates that the leftward end of the SVV genome includes a 665-bp terminal element containing 507 bp of unique sequences bracketed by 79 bp of inverted repeat sequences. These results update our earlier report that an 8-bp inverted repeat sequence flanks the SVV UL segment (Mahalingam et al., 2000). In our previous report, we used PCR amplification of DNA extracted from SVV-infected cells in culture using oligonucleotide primers located at the leftward and the rightward end of the SVV genome. Based on the results presented in the current report, the two oligonucleotides are located on either side of the terminal element. We conclude that PCR amplification of a DNA segment containing a hairpin with a 79-bp stem and a 507-bp loop resulted in the deletion of all but 8-bp of the stem region. Our current sequence of the leftward end initially determined using the recombinant cosmid representing the leftward end of the SVV genome (Fig. 2) is validated by (i) identification of identical sequences in DNA obtained from Vero cells infected with the DHV strain of SVV (Fig. 3); (ii) the detection of these sequences at the leftward end of the genomes of three additional SVV strains (Fig. 5); (iii) the PCR amplification and detection of portions of these sequences in tissues from SVV-infected monkeys (Fig. 6); and (iv) the detection of similar terminal elements at the leftward ends of genomes of other animal varicelloviruses including EHV-1, EHV-4 (Telford et al., 1992, 1998) and PRV (Klupp et al., 2004).

The EHV-1, EHV-4 and PRV genomes have 87, 86 and 82 bp of inverted repeats flanking 945, 668 and 358 bp of unique sequences, at the leftward ends of their respective genomes. The inverted repeat sequences located at the leftward ends of EHV-1, EHV-4, PRV, SVV and VZV contain areas of sequence identity (Fig. 7). These sequences may be important in the cleavage and packaging of the virus genome during replication.
Similar to our observations in SVV (Fig. 1), part of the inverted repeat sequences (32 and 27 bp in EHV-1 and EHV-4, respectively) have also been shown to be present at the junction of the long and short segments of other varicellovirus genomes (Telford et al., 1992, 1998). This feature seems to be absent in the PRV genome, although vestigial inverted repeat sequences flanking the unique long segment have been reported (Demarchi et al., 1990). VZV genome does not have a terminal element that is seen in SVV, EHV-1, EHV-4 and PRV.

The Pac2 site and the GC motif are seen at the leftward end of all five viruses (Fig. 7). Similar sequences have been shown to be important for the cleavage and packaging of bovine herpes virus 1 genomes into virions (Schyns et al., 2003). We observed that the DNA segment located at the leftward end of SVV genome undergoes inversion in infected cells in culture (Fig. 4). Such inversion of the terminal elements located at the leftward ends in EHV-1, EHV-4 or PRV has not been reported. Inverted repeat sequences flanking the entire long segment in different varicelloviruses have been predicted to mediate low level inversion of this region (Davison, 1984; Demarchi et al., 1990). The significance of the inversion of the terminal element in SVV is unclear.

We have demonstrated variation in the size of the unique sequence of the left end terminal element among the genomes of four SVV strains isolated from distinct simian varicella epizootics (Fig. 5). Two of these four SVV strains (MLMV and LVV) contained a duplication of a 46-bp segment located within the unique sequences of genomic DNA isolated from SVV-infected Vero cells. DNA derived from tissues of MLMV-, LVV- and HAZV-infected animals were not available for analysis. Outside of this 46-bp repetition, the terminal element sequences were identical between the four strains. The significance of the presence of direct repeats in some stains of SVV but not in others is unclear, although the presence of direct repeats of different sizes have been reported in EHV-1, -4 and PRV (Telford et al., 1992, 1998; Klupp et al., 2004).

The unique sequences within the terminal elements of EHV-1, EHV-4 and PRV have been shown to contain transcriptional regulatory sequences such as HSV UL9 and nuclear factor-1 (Kupershmidt et al., 1991). We identified an open reading frame that matched with the transcript within the newly detected SVV sequences (Fig. 8). No transcripts or protein coding sequences have been identified with the terminal element of EHV-1, EHV-4 or PRV. Although we identified an open reading frame that matched with the location of the detected transcript (data not shown), it is yet to be determined if it is translated in virus-infected cells.

Taken together, we have identified the leftward end of the SVV genome and found it to be very similar in structure to other animal herpesviruses. These structures may play a role in the replication and packaging of virus DNA during replication.

Materials and methods

Virus, cells and isolation of SVV DNA

SVV strains Delta herpesvirus (DHV), Hazelton herpesvirus (HAZV), Liverpool vervet virus (LVV) and Medical Lake macaque virus (MLMV) were propagated by co-cultivation of African green kidney (Veros or BSC-1) cells with SVV-infected cells as described (Gray and Gusick, 1996). The origin and epidemiology of these SVV isolates have been previously described (Gray, 2004). SVV DNA was extracted from viral nucleocapsids from SVV-infected Vero cells as described (Gray et al., 1992).

Cloning and sequencing of DNA fragment containing the leftward end of the SVV genome

SVV cosmids were generated as previously described (Gray and Mahalingam, 2005). Briefly, SVV DNA was randomly sheared and the ends were repaired and ligated into the Smal site of pWEB: TNC cosmid vector (Epicentre Biotechnologies) and packaged into lambda bacteriophage using MaxPlax packaging extracts (Epicentre Biotechnologies). Recombinant cosmid A containing the SVV left end was initially identified by restriction endonuclease analysis. The termini of the SVV DNA insert within the recombinant cosmid A were identified by DNA sequence analysis using oligonucleotide primers pWEBP1 (5’-TTCATA-CA CGTGCCGTGACTGCCGTT-3’) and pWEBP2 (5’-GA AAAGTGGC CACCTGACGTCTAAG-3’) located on the vector pWEB: TNC.

SVV infection of monkeys and extraction of DNA from infected tissues

SVV infections and procedures involving non-human primates were performed at Tulane National Primate Research Center following appropriate guidelines and protocols approved by the Tulane Institutional Animal Care and Use Committee. Cynomolgous monkeys were infected with SVV either by intratracheal inoculation of $10^3$ plaque forming units of SVV or by natural exposure to cage mates that were inoculated with SVV (Mahalingam et al., 2002). The infected monkeys were euthanized at 14 and 57 days post-intratracheal inoculation and 85 days after natural exposure. Blood, lung and ganglia were obtained at necropsy. Blood mononuclear cells (MNCs) were isolated as described (White et al., 2002a). Total DNA was extracted from MNCs and other tissues as described (White et al., 2002b).

PCR and RT-PCR

Nested-PCR amplification of DNA using primers described in Fig. 2 was performed and products analyzed using radioactively labeled internal oligonucleotides as described (White et al., 2002a,b).

Total cell RNA was isolated from SVV-infected Vero cells using the TRI Reagent protocol (Molecular Research Center Inc., Cincinnati, OH) and treated with DNase I using a commercial kit (Ambion Inc, Austin, TX). RT-PCR was conducted using the Access RT-PCR system (Promega Corp. Madison, WI). RT
Cloning and sequencing of PCR fragments

PCR fragments generated from different SVV strains were cloned using a TA cloning kit (Promega) and selected clones were used for sequence analysis.

The DNA sequence data were analyzed using Vector NTI Suite software (Invitrogen, Carlsbad, CA). The SVV sequences are deposited in the GenBank database under accession number AF275348 (revised in February 2007).

Acknowledgments

This work was supported in part by Public Health Service grants NS32623 (RM) and AI052373 (WLG) from the National Institutes of Health.

We thank Dr. Donald Gilden for useful discussions and Kara Davis for excellent technical assistance. We thank Dr. Vicki Traina-Dorge and the staff at the Tulane National Primate Center for providing tissues from SVV-infected animals. DNA samples were sequenced by the University of Colorado Cancer DNA sequencing and Analysis Core Facility, supported by the NIH/NCI Cancer Center Support Grant CA46934 and also DNA Sequencing Core Facilities of the University of Arkansas for Medical Sciences.

References


Gray, W.L., Gusick, N.J., 1996. Viral isolates derived from simian varicella epizootics are genetically related but are distinct from other primate herpesviruses. Virology 224 (1), 161–166.


