Molecular cloning of integrated simian sarcoma virus: Genome organization of infectious DNA clones
(recombinant DNA/primate retrovirus/restriction enzyme and R-loop analysis/transformation)

KEITH C. ROBBINS, SUSHILKUMAR G. DEVARE, AND STUART A. AARONSON

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Bethesda, Maryland 20205

Communicated by Robert J. Huebner, February 17, 1981

ABSTRACT The integrated form of simian sarcoma virus (SSV) was molecularly cloned in the Charon 16A strain of bacteriophage A. In transfection analysis, the recombinant viral DNAs demonstrated the ability to transform cells in tissue culture at high efficiency. Such transformants possessed typical SSV morphology, expressed simian sarcoma associated virus (SSAV) gag gene products in the absence of virus release, and released SSV after superinfection with a type C helper virus. A physical map of the 5.8-kilobase-pair (kbp) recombinant viral DNA clone, deduced from restriction endonuclease analysis, revealed a 5.1-kbp SSV genome containing 0.55-kbp-long terminal repeats flanked by 0.45 and 0.25 kbp of contiguous host cell sequences. By R-loop analysis, the viral DNA molecule contained two regions of homology to SSV, separated by a 1.0-kbp nonhomologous region. This SSV-specific sequence was shown to be uniquely represented within the normal cellular DNA of diverse mammalian species, including human. Our results demonstrate that this primate transforming retrovirus arose in nature by recombination of a type C helper virus and a host cellular gene.

Transforming retroviruses have been isolated from a number of avian and mammalian species. These viruses appear to have arisen in nature by recombination between a type C helper virus and host cellular information. The latter has been shown to be essential for viral transforming functions (for recent reviews, see refs. 1 and 2). Thus, accumulating evidence indicates that transforming retroviruses may provide an important means of studying cellular genes involved in malignant transformation. The recent application of recombinant DNA techniques to the study of retroviruses has made it possible to isolate and to amplify the genetic information of such viruses for detailed biological and molecular analyses.

The only known transforming retrovirus of primate origin was isolated from a naturally occurring tumor of a woolly monkey (3). This virus, designated simian sarcoma virus (SSV), has been difficult to characterize because it is defective for replication (4), and SSV stocks contain a high excess of an associated type C RNA helper virus, SSAV (4, 5). In an effort to investigate the molecular organization and transforming functions of the SSV genome, we undertook the cloning of SSV proviral DNA from its integrated state within SSV-transformed nonproducer cells.

MATERIALS AND METHODS

Cells and Viruses. Continuous NIH/3T3 (6) and NRK (7) cell lines have been described. Human embryonic lung and woolly monkey skin fibroblasts were established in tissue culture from explants. The isolation of clonal NRK cell lines nonproductively transformed by different SSV variants has been reported (8).

The integrated form of simian sarcoma virus (SSV) was molecularly cloned in the Charon 16A strain of bacteriophage A. In transfection analysis, the recombinant viral DNAs demonstrated the ability to transform cells in tissue culture at high efficiency. Such transformants possessed typical SSV morphology, expressed simian sarcoma associated virus (SSAV) gag gene products in the absence of virus release, and released SSV after superinfection with a type C helper virus. A physical map of the 5.8-kilobase-pair (kbp) recombinant viral DNA clone, deduced from restriction endonuclease analysis, revealed a 5.1-kbp SSV genome containing 0.55-kbp-long terminal repeats flanked by 0.45 and 0.25 kbp of contiguous host cell sequences. By R-loop analysis, the viral DNA molecule contained two regions of homology to SSV, separated by a 1.0-kbp nonhomologous region. This SSV-specific sequence was shown to be uniquely represented within the normal cellular DNA of diverse mammalian species, including human. Our results demonstrate that this primate transforming retrovirus arose in nature by recombination of a type C helper virus and a host cellular gene.

Transforming retroviruses have been isolated from a number of avian and mammalian species. These viruses appear to have arisen in nature by recombination between a type C helper virus and host cellular information. The latter has been shown to be essential for viral transforming functions (for recent reviews, see refs. 1 and 2). Thus, accumulating evidence indicates that transforming retroviruses may provide an important means of studying cellular genes involved in malignant transformation. The recent application of recombinant DNA techniques to the study of retroviruses has made it possible to isolate and to amplify the genetic information of such viruses for detailed biological and molecular analyses.

The only known transforming retrovirus of primate origin was isolated from a naturally occurring tumor of a woolly monkey (3). This virus, designated simian sarcoma virus (SSV), has been difficult to characterize because it is defective for replication (4), and SSV stocks contain a high excess of an associated type C RNA helper virus, SSAV (4, 5). In an effort to investigate the molecular organization and transforming functions of the SSV genome, we undertook the cloning of SSV proviral DNA from its integrated state within SSV-transformed nonproducer cells.

MATERIALS AND METHODS

Cells and Viruses. Continuous NIH/3T3 (6) and NRK (7) cell lines have been described. Human embryonic lung and woolly monkey skin fibroblasts were established in tissue culture from explants. The isolation of clonal NRK cell lines nonproductively transformed by different SSV variants has been reported (8).

One such line, which codes for expression of SSSV gag gene products p12 and p30, was initially referred to as WSV clone 11 (8). This line is designated SSV-11 NRK in the present study. Clonal isolates of simian sarcoma associated virus (SSAV) (4) and mouse amphotrophic virus (AP 129) (9) were also utilized.

Enzymes. Avian myeloblastosis virus DNA polymerase was obtained from the Research Resources Area (Division of Cancer Cause and Prevention, National Cancer Institute). Restriction endonucleases were obtained from New England Biolabs. T4 ligase was purchased from Bethesda Research Laboratories (Rockville, MD). T4 polynucleotide kinase and bacterial alkaline phosphatase came from P-L Biochemicals. Reaction conditions were as recommended by the supplier.

Electrophoresis and Molecular Hybridization. Endonuclease-treated DNA was electrophoresed in 0.6–1.2% agarose gels and subjected to Southern blotting analysis as described (10). Fragments containing SSV DNA were identified by hybridization to 32P-labeled DNA probes. SSSV DNA was prepared by reverse transcription of poly(A)-selected 35S SSSV RNA primed with calf thymus DNA oligonucleotides as described (11). Purified DNA fragments were labeled by the nick-translation method (12). Stringent (13) and relaxed (14) hybridization conditions were used with SSSV and Moloney murine sarcoma virus (MuSV) DNA probes, respectively. DNA fragments detected by molecular hybridization were visualized by autoradiography.

Molecular Cloning. The Charon 16A strain of A phage was propagated in Escherichia coli K-12 DP50topF (15). DNA was purified from CsCl-banded phage as described (16). High molecular weight DNA was extracted from SSV-11 NRK cells by the method of Blin and Stafford (17). EcoRI fragments containing SSV DNA were partially purified by preparative agarose gel electrophoresis. EcoRI-cleaved vector and partially purified SSV DNA were mixed with T4 ligase and packaged in vitro into phage particles as described (18). In vitro-packaged recombinant DNA demonstrated 3.5 × 10⁶ plaque-forming units/μg of DNA as compared to 1 × 10⁵ for EcoRI-cleaved and 2 × 10⁵ for untreated Charon 16A DNA. Plaques containing SSV DNA were identified by the method of Benton and Davis (19) with SSSV cDNA as probe. All work involving recombinant phage was performed in a P2 containment facility in accordance with National Institutes of Health guidelines for recombinant DNA research.

Electron Microscopy. For R-loop analysis (20), A-SSSV-11 CI DNA (3 μg/ml) was incubated with SSSV 70S RNA (6 μg/ml) for 16 hr at 52°C in hybridization buffer [70% formamide/0.1 M N-tris(hydroxymethyl)methylglycine (Tricine)-NaOH, pH 8.0/0.5 M NaCl/10 mM EDTA]. The sample was diluted 1:10

Abbreviations: SSV, simian sarcoma virus; SSSV, simian sarcoma associated virus; MuSV, murine sarcoma virus; LTR, long terminal repeat; kbp, kilobase pair(s).
in hybridization buffer containing cytochrome at 50 µg/ml and an aliquot was spread onto a hypophase of H2O. The nucleic acids were transferred to Parlodion-coated grids, stained with 0.1 mM uranyl acetate in 90% ethanol, and shadowed with platinum/palladium. Nucleic acid contour lengths were measured with the aid of a Tektronix computerized graphic system. The dX174 DNA and its replicative form were used as molecular weight standards.

Radioimmunoassays. The M, 12,000 (p12) and 30,000 (p30) structural proteins of SSV were isolated as described (8). Each was labeled with ¹²⁵I at high specific activity (10 µCi/µg; 1 Ci = 3.7 x 10¹² becquerels) by the chloramine-T method of Greenwood et al. (21). Competition radioimmunoassays were performed as reported (22).

DNA Transfection. Transfection of NIH/3T3 cells with molecularly cloned or cellular DNA was performed by the calcium phosphate precipitation technique (23) as modified by Wigler et al. (24). Briefly, 1 ml of 0.25 M CaCl₂ containing varying amounts of test DNA and 50 µg of salmon sperm DNA as carrier was mixed with an equal volume of a 50 mM Heps, pH 7.1/280 mM NaCl/1.5 mM sodium phosphate. The calcium phosphate precipitate was added to a 10-cm Petri dish, in which 1–2 x 10⁶ NIH/3T3 cells had been plated 24 hr earlier. After overnight incubation, the supernatant was removed, and 10 ml of Dulbecco’s modified Eagle’s medium supplemented with 5% calf serum was added. Transformed foci were scored at 14–21 days.

For analysis of SSV rescue, individual transformed foci induced at limiting DNA concentration were picked by the cloning cylinder technique, grown to mass culture, and superinfected with mouse amphotropic type C virus, transforming virus that induced foci with morphologic features characteristic of SSV was rescued. All of these results suggested that EcoRI was suitable for cloning intact SSV from SSV-11 NRK cellular DNA.

RESULTS

Selection of Restriction Endonuclease Suitable for Molecular Cloning of Integrated SSV. In order to achieve molecular cloning of infectious SSV provirus in its integrated form, we attempted to find a restriction enzyme compatible with available vectors that would not cleave SSV proviral DNA. DNA prepared from SSV-11 NRK nonproducer cells was treated with restriction enzymes including BamHI, HindIII, Sal I, and EcoRI and subjected to Southern blotting analysis using SSV cDNA as a probe. Cleavage by EcoRI resulted in a single 5.8-kilobase-pair (kb) DNA band; each of the other enzymes tested yielded more than one band (data not shown).

We next compared the transforming activities of EcoRI-cleaved and untreated SSV-11 NRK DNA by transfection analysis. Their focus-forming activities on NIH/3T3 cells were found to be similar and ranged from 3 to 10 focus-forming units/100 µg of cellular DNA. Moreover, when transformed foci induced by EcoRI-cleaved SSV-11 NRK DNA were superinfected with mouse amphotropic type C virus, transforming virus that induced foci with morphologic features characteristic of SSV was rescued. All of these results suggested that EcoRI was suitable for cloning intact SSV from SSV-11 NRK cellular DNA.

Molecular Cloning of the Integrated SSV Genome. EcoRI-cleaved SSV-11 NRK DNA was subjected to agarose gel electrophoresis in an effort to enrich for SSV proviral sequences. The 5.8-kbp DNA fraction was ligated to EcoRI-cleaved DNA purified from the Charon 16A strain of λ phage (15), packaged in vitro into phage particles as described (18), and plated onto E. coli K-12 DP50supF. The resulting plaques were tested for reactivity with SSV cDNA by the method of Benton and Davis (19). Of 1.7 x 10⁵ plaques analyzed, 2 were scored as positive. Each was plaque-purified twice prior to further analysis. DNA extracted from the purified phage clones, designated λ-SSV-11 Cl1 and Cl2, was treated with EcoRI and subjected to Southern blotting analysis. In each case, EcoRI-cleavage yielded a 5.8-kbp insert which was readily detected by hybridization with SSV cDNA (Fig. 1).

Biological Activity of Cloned SSV DNA. To determine whether the recombinant DNA clones possessed biological activity, they were analyzed by transfection on NIH/3T3 cells. λ-SSV-11 Cl1 and Cl2 DNAs each demonstrated high-titered focus-forming activity that was not significantly altered by EcoRI cleavage (Table 1). Focus formation was a linear function of amount of DNA added, indicating that a single DNA molecule was able to induce a transformed focus. When individual transformed foci were selected by the cloning cylinder technique, grown to mass culture, and superinfected with amphotropic mouse type C helper virus, focus-forming activity characteristic of SSV was rescued from each transformant tested. Previous studies have shown that SSV-11 NRK cells express SSV gag gene encoded p12 and p30 (8). We assayed transformants induced by SSV recombinant DNA clones for expression of these

<table>
<thead>
<tr>
<th>Table 1. Biologic activity of λ-SSV-11 recombinant DNA clones</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Recombinant DNA</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>λ-SSV-11 Cl1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>λ-SSV-11 Cl2</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* DNA transfection was performed with 50 µg of salmon sperm DNA as carrier. Focus formation on NIH/3T3 cells was scored at 14–21 days. DNA added is shown as µg/plate.

† Shown as focus-forming units/pmol of viral DNA.

² Individual transformed foci were isolated at limiting DNA concentration, grown to mass culture, and superinfected with a mouse amphotropic virus. At 2–3 weeks, tissue culture fluids were assayed for focus formation on NIH/3T3 cells.
proteins and found that they expressed SSAV p12 and p30 at levels ranging from 30 to 120 ng/mg of cell protein (data not shown). Thus, the phenotype of transformants induced by the SSV recombinant DNA clones was indistinguishable from that of transformants induced by the parental virus.

Physical Map of λ-SSV-11 CLI DNA. To construct a physical map of λ-SSV-11 CLI DNA, restriction enzyme cleavage products were electrophoresed on agarose gels and visualized by ethidium bromide staining. The relative location of cleavage sites on a linear map was determined by appropriate double digestions. The results of these studies are summarized in Fig. 2.

Proviral DNAs of avian and mammalian retroviruses so far examined have been shown to contain long terminal repeats (LTRs) (27–30). We attempted to localize analogous regions on the λ-SSV-11 CLI physical map by determining whether a specific constellation of restriction enzyme sites was repeated within the molecule. Ps I and Kpn I cleaved the molecule in identical locations with respect to two Sae I sites located near either terminus at approximately 0.8 and 5.5 kbp on the restriction map. It seemed possible that these regions, each approximately 0.5 kbp long, might correspond to the LTRs of SSV.

In order to establish more firmly the presence of LTRs and define their length, we constructed a detailed restriction map of the terminal 1.3- and 1.0-kbp Xba I fragments of λ-SSV-11 CLI DNA. HindIII and Alu I were chosen for this purpose because the sites at which they cleave occur frequently in DNA. An identical constellation of two HindIII and four Alu I sites was observed in each fragment within the regions defined by Ps I and Kpn I sites. Several unique HindIII and Alu I sites were detected outside of these domains. Results demonstrated that terminal redundancies were present in SSV-11 DNA and defined their length as 0.55 ± 0.02 kbp. From the location of the two LTRs within λ-SSV-11 CLI DNA, it was possible to establish the length of SSV genomic DNA as 5.1 kbp and the presence of flanking NRK cellular DNA sequences extending 0.45 and 0.25 kbp to the left and right ends, respectively, of the proviral DNA.

Orientation of λ-SSV-11 CLI DNA with Respect to SSV RNA. It is known that the viral gag gene resides within the 5'-region of the type C viral RNA genome (for review, see ref. 31). The fact that SSV-11 codes for SSV gag gene products made it possible to devise a means of assigning these coding sequences to a specific region of the λ-SSV-11 CLI DNA physical map. We took advantage of knowledge that mammalian type C viruses are evolutionarily related and of the availability of a molecular clone of Moloney MuSV DNA (32). Moloney MuSV DNA subclones composed of murine leukemia virus LTR or gag gene sequences were nick-translated and hybridized to Southern blots of Xba I-cleaved λ-SSV-11 CLI DNA. The four Xba I fragments (Fig. 3, lane 2) were ordered 1.3, 1.6, 1.9, and 1.0 kbp from left to right on the λ-SSV-11 CLI physical map (Fig. 2). As shown in lane 3 of Fig. 3, the LTR subclone of Moloney MuSV DNA detected only the 1.3- and 1.0-kbp Xba I fragments which were shown above (Fig. 2) to contain the SSV LTRs. In contrast, the Moloney MuSV gag gene probe hybridized with the 1.3- and 1.6- but not the 1.9- or 1.0-kbp fragments (Fig. 3, lane 4). These results localized SSV gag gene-related information at the left end of the physical map of λ-SSV-11 CLI DNA and thus oriented this end as 5' with respect to SSV genomic RNA.

Detection and Localization of SSV Nucleotide Sequences Not Found in SSV. The transforming retroviruses studied so far have been shown to contain nucleotide sequences unrelated to those of a type C helper virus in addition to helper virus-related information (for recent reviews, see refs. 1 and 2). In an effort to detect and to localize such sequences in λ-SSV-11 CLI DNA, we performed R-loop analysis of this clone with SSV genomic RNA. A representative of 30 R-loop structures analyzed is shown in Fig. 4. Two regions of DNA-RNA homology, 3.2 ± 0.2 and 0.7 ± 0.01 kbp, were detected. Within the larger region of homology, two loops of single-stranded RNA representing SSV sequences deleted from the SSV genome were observed. Of particular note was the finding that the two regions of DNA-RNA homology were separated by 1.0 ± 0.1 kbp of λ-SSV-11 CLI DNA that lacked homology with SSV RNA. These results indicated that λ-SSV-11 CLI contained a contiguous 1.0-kbp region that was unrelated to helper virus and was located within the transforming virus genome.

In order to orient the R-loop structure with respect to SSV genomic RNA and the λ-SSV-11 CLI DNA physical map, we took advantage of the fact that SSV codes for a gag gene product containing SSV p12 and p30. Only the larger region of DNA-RNA homology possessed sufficient coding capacity for this gene product. Thus, the 3.2- and 0.7-kbp regions of DNA-RNA homology must be located 5' and 3', respectively, on SSV genomic RNA.

As an independent approach toward localization of SSV unrelated sequences within λ-SSV-11 CLI DNA, we utilized the Southern blotting technique. From the results of R-loop analysis, the region located at approximately 4.0–5.0 kbp on the λ-SSV-11 CLI DNA physical map (Fig. 2) should contain information unrelated to SSV. To test this possibility, a 1.6-kbp Sac I DNA fragment located 4.0–5.6 kbp on the SSV-11 CLI DNA physical map was purified. After cleavage with Bgl II, Bgl I, or XbaI, the resulting fragments were subjected to Southern blotting analysis and hybridization with a nick-translated 8.8-kbp SSV DNA that had been cloned from the circular form of the unintegrated provirus (F. R. Anderson, personal communication).

**Fig. 2.** Restriction map of λ-SSV-11 CLI DNA. Sites of restriction enzyme cleavage within the DNA insert were determined by double digestion analysis. The sites of HindIII and Alu I cleavage within the terminal Xba I DNA fragments were determined by using the partial digestion method of Smith and Birnstiel (26). Wavy lines, NRK cellular flanking sequences; LTR, long terminal repeat.
DISCUSSION

The recent application of recombinant DNA techniques to the study of retroviruses has made these viruses much more amenable to detailed biochemical analysis. Our strategy for molecular cloning of the primate transforming retrovirus SSV involved isolation of a DNA fragment containing intact, linear SSV DNA and host flanking sequences from cells nonproductively transformed by a clonal SSV variant. This approach required first the selection of a restriction enzyme, EcoRI, that did not cleave within the integrated viral genome. We recovered one SSV recombinant DNA phage per 10^5 plaque-forming units

**Fig. 5.** Localization of nucleotide sequences unrelated to SSV on the restriction map of λ-SSV-11 C11 DNA. A DNA fragment located at 4.0–5.6 kbp on the λ-SSV-11 C11 restriction map was purified by agarose gel electrophoresis after Sac I cleavage of λ-SSV-11 C11 DNA. The purified 1.6-kbp DNA fragment (lane 1) as well as DNA products resulting from cleavage with Bgl II (lane 2), Bgl I (lane 3), and Xba I (lane 4) were visualized by ethidium bromide staining. DNA fragments containing SSV-related nucleotide sequences were identified by Southern blotting and hybridization with nick-translated, molecularly cloned SSV genomic DNA as probe (lanes 5–8, respectively). Fragments of HindIII-digested λ and Hae III-digested φX174 replicative form DNAs served as size standards (arrows from top to bottom): 1.6, 1.3, 1.05, 0.95, 0.72, 0.61, and 0.29 kbp.

**Fig. 6.** Detection of homology between normal cellular DNA and SSV nucleotide sequences not found in SSV. Cellular DNAs were treated with EcoRI, fractionated by agarose gel electrophoresis, and subjected to Southern blotting analysis using, as a hybridization probe, the SSV-specific DNA fragment derived from 4.0–5.0 kbp on the restriction map of λ-SSV-11 C11 DNA. Cellular DNAs included those of SSV-11 NRK (lane 1), uninfected NRK (lane 2), woolly monkey fibroblasts (lane 3), and human embryo lung fibroblasts (lane 4). Size markers (arrows from top to bottom): 23.0, 15.0, and 5.8 kbp.
after enrichment for viral sequences by a single cycle of preparative agarose gel electrophoresis. Whereas additional purification procedures can enrich further for a particular gene (33-35), our results demonstrate that extensive purification is not required for successful isolation of single copy genes.

Transfection analysis revealed that the recombinant DNA clones transformed mouse cells in tissue culture at a high efficiency. The morphologic characteristics of DNA transfectants as well as foci induced by sarcoma virus rescued from such transfectants were indistinguishable from those of foci induced by SSV. The λ-SSV-11 recombinant DNA clones were also shown to code for expression of SSAV gag gene products p12 and p30, additional markers of the parental SSV-11 genome (8). Thus, the λ-SSV-11 recombinant DNA clones appeared to possess all of the known biological functions of SSV.

The molecular organization of the integrated SSV provirus in λ-SSV-11 C1 I DNA, as deduced from restriction mapping and R-loop analysis, demonstrated several important features (Fig. 7).LTRs approximately 0.55 kbp in length were defined by the presence of an identical constellation of restriction sites located near either end of λ-SSV-11 C1 I DNA. Analogous structures at the termini of avian and murine retroviruses have been shown to contain putative regulatory signals for initiation and termination of transcription (36-38). Nucleotide sequence analysis of retroviral LTRs has also indicated their structural similarities with transposable elements, suggesting that the LTRs may play an important role in proviral DNA integration into the cellular genome.

R-loop analysis helped to explain the basis for the replication-defective nature of SSV. From knowledge of the type C viral gene order as 5' gag-pol-env 3' (31), the 1.5-kb deletion of SSAV sequences toward the 3' end of the viral RNA was likely to lie within the SSAV env gene. Of the two SSAV deletion loops detected within the major 3.2-kbp region of SSV/SSAV homology, the larger likely corresponded to the region of the SSV pol gene. Although the smaller deletion (0.2 kbp) appeared to be located at or near the beginning of the viral gag gene, this deletion did not impair the ability of the SSAV recombinant DNA clones to synthesize SSAV gag gene products. Precise localization of SSAV coding sequences deleted from SSAV will require detailed nucleotide sequence comparison of molecularly cloned SSAV with our SSAV recombinant DNA clones.

The 5.1-kbp SSAV DNA genome was shown to be composed of SSAV homologous regions, 3.2 and 0.7 kbp, respectively, at 5' and 3' ends of the molecule connected by a continuous 1.05-kbp region of SSAV nonhomologous information. Nucleotide sequences related to the SSAV-specific region were detected at single copy number within normal cellular DNAs of several mammalian species. These findings strongly argue that SSAV arose by recombination of SSAV with a well-conserved cellular gene. Because the cloned SSAV-specific sequence readily detected related information in human cellular DNA, it should be possible to determine whether altered or enhanced expression of SSAV related sequences is involved in naturally occurring human malignancies.

![Fig. 7. Molecular organization of integrated SSAV DNA in SSAV-11 NRK cellular DNA. Wavy lines represent flanking NRK cellular DNA sequences. Open boxes represent LTRs. Darkened rectangle represents SSAV nucleotide sequences not found in SSAV.](image-url)