Comparative Restriction Endonuclease Maps of Proviral DNA of the Primate Type C Simian Sarcoma-Associated Virus and Gibbon Ape Leukemia Virus Group

CECELIA D. TRAINOR, FLOSSIE WONG-SТАAL, AND MARVIN S. REITZ, JR.*
Division of Cancer Treatment, Laboratory of Tumor Cell Biology, National Cancer Institute, Bethesda, Maryland 20205

Received 10 July 1981/Accepted 5 October 1981

Extrachromosomal DNA was purified from canine thymus cells acutely infected with different strains of infectious primate Type C viruses of the woolly monkey (simian) sarcoma helper virus and gibbon ape leukemia virus group. All DNA preparations contained linear proviral molecules of 9.1 to 9.2 kilobases, at least some of which represent complete infectious proviral DNA. Cells infected with a replication-defective fibroblast-transforming sarcoma virus and its helper, a replication-competent nontransforming helper virus, also contained a 6.6- to 6.7-kilobase DNA. These proviral DNA molecules were digested with different restriction endonucleases, and the resultant fragments were oriented to the viral RNA by a combination of partial digestions, codigestion with more than one endonuclease, digestion of integrated proviral DNA, and hybridization with 3' and 5'-specific viral probes. The 3' and 5'-specific probes each hybridized to fragments from both ends of proviral DNA, indicating that, in common with those of other retroviruses, these proviruses contain a large terminal redundancy at both ends, each of which consists of sequences derived from both the 3' and 5' regions of the viral RNA. The proviral sequences are organized 3',5'-unique-3',5'. Four restriction enzymes (KpnI, Smal, PstI, and SstI) recognized sites within the large terminal redundancies, and these sites were conserved within all the isolates tested. This suggests that both the 3' and 5' ends of the genomic RNA of these viruses are extremely closely related. In contrast, the restriction sites within the unique portion of the provirus were not strongly conserved within this group of viruses, even though they were related along most of their genomes. Whereas the 5' 60 to 70% of the RNA of these viruses was more closely related by liquid hybridization experiments than was the 3' 30 to 40%, restriction sites within this region were not preferentially conserved, suggesting that small sequence differences or point mutations or both exist throughout the entire unique portion of the genome among these viruses.

Horizontally transmitted type C retroviruses have been isolated many times from subhuman primate tissues. They have been isolated from two different species of primates and can be distinguished from one another by immunological (16) and nucleic acid (22) analyses. Nonetheless, all are very closely related and appear to form a single class of virus (4). The first isolate, simian sarcoma virus, was obtained from a fibrosarcoma of a woolly monkey (Lagothrix spp.) (32, 35). Similar viruses were subsequently isolated from a second primate species, the white-handed gibbon (Hylobates lar). Among these viruses are isolates from gibbons with lymphosarcoma (GaLVSP) (13), granulocytic leukemia (GaLVSEATO) (12), and acute lymphocytic leukemia (GaLVH) (3). A similar virus (GaLVH) was also isolated from the brain tissue of gibbon apes which had been inoculated with tissue extracts of a human patient with kuru (33), but which had no evident leukemia. Two of the above four isolates have been shown to induce neoplasia in primates in laboratory experiments. One isolate, simian sarcoma virus, is a mixture of a replication-defective fibroblast-transforming sarcoma virus (SSV) and a replication-competent nontransforming helper virus (SSAV) (34). The SSV-SSV complex induces fibrosarcomas, fibromas, and astrocytomas in marmosets (35). The other isolate, GaLVSEATO, induces chronic myelogenous leukemia in young gibbon apes (14).

The extensive genetic similarity within this family of viruses in conjunction with their apparently different biological properties makes them an attractive system for correlating the diver-
gence of nucleic acid sequences with differences in pathology. Moreover, the lack of endogenous DNA sequences related to these viruses in primates (4) permits an easier analysis of proviral sequences than with viruses which have extensive, related endogenous sequences in host DNA.

Previous results localized the most divergent regions of GaLV<sub>SF</sub>, GaLV<sub>H</sub>, and SSV to the 3' 30 to 40% of the viral genome by liquid molecular hybridization (22). Restriction endonuclease and Southern blot hybridization techniques were employed to construct comparative maps of the genomes of SSV and GaLV and to further localize the divergent and conserved regions of the genomes. Restriction maps of unintegrated and integrated proviral DNA were obtained for SSV, GaLV<sub>SEATO</sub>, and GaLV<sub>Ba</sub>. These maps were oriented 3' to 5' with respect to the viral RNA genome by the use of labeled probes specific for the 3' and the 5' ends of the viral genome. Comparative maps of these proviruses are presented here.

**MATERIALS AND METHODS**

**Viruses and cells.** The viruses used in these studies were simian (woolly monkey) sarcoma and helper virus complex SSV, SAV, and the gibbon ape leukemia viruses. GaLV<sub>SEATO</sub> (12), GaLV<sub>SF</sub> (13), GaLV<sub>H</sub> (3), and GaLV<sub>Ba</sub> (33). SSV-SSV was obtained from a culture of marmoset tumor cells (17API) explanted from an animal inoculated with the original virus isolate. SSV is produced in 10-fold excess over SSV in this culture (P. Markham, personal communication). Marmoset fibroblasts (HF) were infected with this virus and were used in some studies (HF/SSV). SSV was grown in the human rhabdomyosarcoma cell line A204. GaLV<sub>SEATO</sub> and GaLV<sub>Ba</sub> were grown in bat lung fibroblasts (CCL-88). GaLV<sub>SF</sub> and GaLV<sub>HB</sub> were grown in the primary tumor lines UCD-144 and 6G-1, respectively.

**Preparation of labeled viral probes.** Viral RNA was prepared from a virus purified by equilibrium banding in sucrose gradients. The virus was diluted to 1% in sodium dodecyl sulfate-1 mg of pronase per ml, incubated at 37°C for 30 min, extracted with PCC-9 (phenol-cresol-chloroform) and precipitated with ethanol-salt. The high-molecular-weight RNA (50 to 70S) was then purified by velocity gradient sedimentation as described previously (3) and labeled with <sup>125</sup>I by the method of Comerford (2) to a specific activity of 5 x 10<sup>7</sup> to 10 x 10<sup>8</sup> cpm/μg.

<sup>32</sup>P-<sub>P</sub>DNA. Purified 70S RNA was transcribed with avian myeloblastosis virus reverse transcriptase. The reaction contained in 0.1 ml: 300 μ of reverse transcriptase; 0.05 M Tris, pH 7.8; 0.05 M dithiothreitol; 6 x 10<sup>-3</sup> M MgCl<sub>2</sub>; 0.06 M KCl; 0.5 x 10<sup>-3</sup> M each dATP, dGTP, and TTP; 10<sup>-3</sup> M [<sup>32</sup>P]-dTTP (400 to 600 Ci/mmole); and 10 mg of randomly digested calf thymus DNA per ml. The cDNA was purified by PCC-9 extraction and by cetyltrimethylammonium bromide and ethanol precipitation as described previously (21).<sup>125</sup>I-labeled RNA representing the 3' 20 to 25% of the viral genome was prepared from SSV, SSV-SSV 50 to 70S RNA by size selection of polyadenylic acid-containing subgenomic fragments as described previously (22). "Strong-stop" [<sup>32</sup>P]-DNA (8), complementary to the 145 5' nucleotides of the 50 to 70S viral RNA, was prepared in endogenous reverse transcription reactions with the following modifications of the conditions described by Lovinger and Schochetman (17). [<sup>32</sup>P]-dCTP was present at 2.5 x 10<sup>-6</sup> M, melittin (50 μg/ml) was substituted for Nonidet P-40 (1), and the incubation period was shortened to 1 to 2 h. Strong-stop cDNA was purified as described for total cDNA (see above) and electrophoresed on an 8% acrylamide-0.27% N,N-methylenebisacrylamide slab gel (0.5 mM) for 3 h at 150 V. The cDNA was visualized by autoradiography, and the prominent band at 140 to 145 nucleotides in size was excised, electroeluted into a dialysis bag (18), and concentrated by ethanol-salt precipitation.

**Preparation of integrated and unintegrated proviral DNA.** High-molecular-weight DNA containing integrated provirus was prepared from chronically infected virus-producing cells. The cells were suspended in TE buffer (10 mM Tris-hydrochloride [pH 7.4], 10 mM EDTA), extracted with PCI-9 (phenol-isooamyl alcohol-chloroform, 1:0.05:1; pH 9) and with chloroform alone, and then treated with 1% sodium dodecyl sulfate and 1 mg of pronase per ml (1 h, 37°C). The extraction procedures were repeated, and the DNA was precipitated with ethanol.

Unintegrated proviral DNA from acutely infected canine thymus cells was purified by the Penman (19) method for preparing cytoplasmic nucleic acids. Cells were harvested 6 to 16 h after infection (multiplicity of infection, 10 infectious units of concentrated virus per cell) and suspended in 5 ml of 10 mM Tris-hydrochloride [pH 7.4], 2.5 mM MgCl<sub>2</sub>-5 mM NaCl to a final concentration of 10<sup>6</sup> cells per ml. One volume of 0.2% Triton was added, and the mixture was incubated at 22°C for 20 min and centrifuged at 1,000 x g for 10 min. The supernatant was removed, the washed supernatant was resuspended in the same buffer. A 2:1 solution of 10% Tween 40 and 10% deoxycholic acid was prepared immediately before use, and one-seventh of a volume was added to the resuspended pellet. This mixture was blended with a Vortex mixer for 30 s and centrifuged as described above. The supernatants were combined, diluted to 1% in sodium dodecyl sulfate and 1 mg of pronase per ml, incubated at 37°C for 1 h, extracted with PCI-9, and then precipitated with ethanol. The proviral DNA was dissolved in 1 ml of 10 mM Tris-hydrochloride (pH 7.4) with 1 mM EDTA mixed with 4 ml of a solution of cesium chloride (1.7 g/ml) and ethidium bromide (0.3 g/ml) and was centrifuged at 43,000 rpm in a Beckman SW50.1 rotor for 65 to 70 h. Fractions were collected and their densities were determined by refractive index. The 1.55-g/ml region of the gradient was pooled and extracted with n-butanol to remove ethidium bromide, and the DNA was concentrated by ethanol precipitation. Samples were assayed for proviral DNA by Southern blot hybridizations (28). Alternatively, some preparations of unintegrated proviral DNA were purified by the Hirt procedure (9) and then were further purified on CsCl as described. Similar results were obtained with both methods.

**Restriction nuclease analysis.** DNA samples were
incubated (37°C, 4 to 6 h) with 1 to 2 U of restriction nuclease (Bethesda Research Laboratories, Frederick, Md.) per μg of DNA under conditions specified by the manufacturer. Reactions were stopped by the addition of one-fifth volume of Biophore (Biorad Laboratories, Richmond, Calif.) containing 10 mM EDTA and 10% sodium dodecyl sulfate.

**Southern blot hybridization.** DNA was electrophoresed on 8-mm slab gels containing 0.8% agarose in 0.04 M Tris-hydrochloride (pH 7.4)–18 mM NaCl–0.02 M sodium acetate–2 mM EDTA (pH 8) at 120 mA for 16 to 20 h. DNA was transferred to nitrocellulose paper and hybridized to viral probes by the method of Southern (28) as modified by Knetter and Kelly (15).

Hybridization to DNA immobilized on the filters was performed in 50% formamide–3× SSC (1× SSC = 0.015 M sodium citrate plus 0.15 M NaCl, pH 7) with 0.5 × 10^6 cpm of 125I-labeled RNA or 32P-labeled DNA as described previously (36), except when 125I-labeled 3' RNA or 32P-labeled strong-stop cDNA was used. Only 4 × 10^6 to 5 × 10^6 cpm of strong-stop cDNA was used per filter, since it represents only a limited region of the genome, whereas 2 × 10^9 to 3 × 10^9 cpm of 125I-labeled 3' RNA was used for the same reason.

**RESULTS**

**Forms of unintegrated provirus.** Unintegrated proviral DNAs from SSV-SSV, GalVSEATO, and GalVBR were analyzed by Southern blot hybridization. Labeled probes from SSV-SSV were used for all three viruses, since all are related by molecular hybridization over most of their genomes (22). In each case, the most prominent band was around 9.15 kilobases (kb). This band (shown for SSAV-SSV in Fig. 1) corresponds to linear proviral DNA and contains infectious DNA as assayed by transfection experiments (our unpublished results with P. Markham). Often one or more fainter bands migrating faster than those of complete linear provirus were also observed. One of these bands migrated slightly ahead of the main band at about 7.8 kb and probably corresponds to a supercoiled circular form of the provirus, as observed for other retroviruses (11, 25–27, 37).

The most intense of the minor bands, shown in Fig. 1, migrated at about 6.6 kb, about the size expected for the sarcoma virus component of SSV-SSV, based on the size of its genomic RNA (24). This band was readily apparent in preparations of SSV-SSV, but was not evident in preparations of SSAV alone or of any of the GaLV isolates, all of which lack an associated sarcoma virus. This band probably represents the linear viral genome of the SSV, although fibroblast transformation upon transfection with this DNA has not been demonstrated. Restriction maps of this fragment are identical with maps generated from two clones of SSV constructed in this laboratory (5). The identity of other minor bands sometimes observed is not known. With the exception of the putative SSV provirus, all of the minor bands were at sufficiently low levels so that they would not interfere with the analysis of main band DNA and would not be considered further.

**Orientation of BamHI and HindIII restriction maps of SSAV to viral genomic RNA.** The fragments generated by digestion with BamHI and HindIII were oriented, with respect to each other, by partial and double digests of unintegrated DNA and, with respect to viral RNA, by hybridization to end-specific probes. Digestion of integrated, proviral DNA was used to confirm the positions of internal fragments as assigned by digestion of unintegrated, linear DNA.

An autoradiogram of a typical preparation of SSAV-SSV unintegrated DNA is shown in Fig. 1. Lane A contained a 9.15-kb linear SSAV

![Figure 1](http://jvi.asm.org/ Downloaded from http://jvi.asm.org/ on November 6, 2017 by guest)
proviral DNA and two very faint, faster-migrating bands, which probably represent a circular SSAV provirus and a linear SSV proviral DNA, as discussed above. The pattern with proviral DNA of SSAV alone was identical (not shown), indicating that not one of the bands in the BamHI digest is due to the SSV component. This result indicates that BamHI cleaved the SSAV proviral DNA at least twice. SSAV-SSV unintegrated DNA was partially digested with BamHI (Fig. 1, lane C). Besides the 4.1-, 3.3-, and 1.55-kb fragments, two additional fragments of 5.8 and 4.5 kb appeared. The lack of a 7.4-kb fragment in this digest indicated that the large fragments are not adjacent, therefore; the fragment order is 4.1, 1.55, and 3.3. This was confirmed by the digestion of integrated proviral DNA with BamHI. Cell-provirus junction fragments from multiple integration sites were heterogeneous in size and did not yield discrete, intense bands. BamHI digests of integrated provirus from both 71AP1 (lane E) and HF/SSAV-SSV cell DNA (not shown) yielded only the 1.55-kb fragment of the three observed with unintegrated provirus, confirming that this is an internal fragment and showing that there is no single dominant integration site in the DNA of either of these infected cell lines.

Fragments generated by a second enzyme, HindIII, are shown in Fig. 1 (lanes D and F). Four fragments (4.25, 1.9, 1.6, and 1.4 kb) were generated from unintegrated SSAV-SSV. The 1.9- and 1.4-kb fragments, but not the 4.25- and 1.6-kb fragments, are internal since they were observed in HindIII digests of infected cell DNA (Fig. 1, lane F). These internal fragments were oriented in double digests (see below).

To orient the DNA fragments generated by BamHI and HindIII 3' to 5' with respect to viral RNA, 125I-labeled RNA, highly enriched for the 3' 20 to 25% of the viral genome (22), was hybridized to SSAV-SSV proviral DNA digested with BamHI or HindIII. The relative intensities of these bands were quantitated by scanning the autoradiograms of the Southern blots with a densitometer (Fig. 2), tracing the peaks onto graph paper, cutting them out, weighing them on an analytical balance, and calculating their respective weights (Table 1). With BamHI digests of unintegrated proviruses, the 3' specific probe hybridized to both the 4.1- and the 3.3-kb end fragments. The 3.3-kb fragment was more intensely labeled, as shown by densitometer tracings (Fig. 2A; Table 1). There was little or no hybridization to the middle (1.55-kb) fragment. Hybridization of an end-specific probe to both end fragments indicated that the SSAV provirus is terminally redundant, as are the proviruses of avian (10, 11, 26), feline (27), and murine (6, 25) retroviruses. The greater intensity of the 3.3-kb fragment indicated that it corresponds to the 3' end of the viral RNA, since the RNA probe was long enough to contain sequences unique to the proviral 3' end as well as to those common to both ends. To rule out the possibility that this result was due to a qualitative variation in the transfer of viral DNA fragments to nitrocellulose, the DNA on the same filter was rehybridized to unfractionated SSAV 125I-labeled RNA. All three BamHI fragments were evident, and the 4.1-kb fragment was the most intense, strengthening our interpretation (Fig. 2C; Table 1). This orients the BamHI restriction map to viral genomic RNA as 5'–4.1, 1.55, 3.3–3'.

The end fragments produced by HindIII digestion of SSAV proviral DNA, as determined above, are 4.25 and 1.6 kb. Both bands were specifically labeled by the 3' 125I-labeled RNA (Fig. 2B; Table 1), again indicating that the SSAV provirus is terminally redundant. However, the intensity of the 4.25-kb band increased significantly relative to the 1.6-kb band when an unfractionated probe was rehybridized to the same filter (Fig. 2D; Table 1), indicating that the 4.25-kb fragment is the 5' end of the genome, and the 1.6-kb fragment is the 3' end. The two putative internal fragments were not significantly labeled by this probe, but appeared upon subsequent rehybridization with the total probe, confirming their internal position within the provirus. The order of the two internal HindIII fragments was determined through double digests of integrated SSAV proviral DNA with BamHI and HindIII (data not shown). Internal bands of 1.9 and 1.2 kb were found. This indicated that the HindIII map is 5'–4.25, 1.4, 1.9, 1.6–3', because if the internal bands were in the opposite orientation, the 1.9-kb fragment would be cleaved by BamHI.

The BamHI and HindIII maps of SSAV were used as references for ordering fragments produced by other restriction enzymes. Double digests with BamHI or HindIII were used in combination with complete and partial digests to construct maps for other enzymes, as described below.

Restriction endonucleases which cleave within the proviral LTR sequences of SSAV and GaLV. Restriction endonucleases which cleave within the large terminal redundancy (LTR) regions of the provirus are useful for several reasons. Since the sizes of the internal fragments cleaved from integrated DNA add up to a complete provirus minus one LTR sequence, they define both the size of the LTR region and that of the unique portion of the provirus. They cleave the entire unique region of the integrated provirus from host flanking sequences, facilitating the mapping of integrated DNA. The nucleotide sequence for 5' (strong-stop) cDNA of GaLVH has been de-
FIG. 2. Hybridization of BamHI and HindIII digestion fragments of SSAV-SSV unintegrated proviral DNA to 3' and total 125I-labeled RNA. The DNA from SSAV-SSV was digested with BamHI (panels A and C) or HindIII (panels B and D) and was electrophoresed on 0.8% agarose as described in the text. After being transferred to nitrocellulose, the fragments were hybridized to 125I-labeled RNA representing the 3' approximate 20% of the viral genome (panels A and B). Bands of hybridization were visualized by autoradiography with Kodak XR2 X-ray film. The autoradiographs were developed and scanned with a Transidyne 2995 scanning densitometer with white light. The same filters were later hybridized to total 125I-labeled RNA (panels C and D) after the filters were washed sufficiently free of radioactivity as to not obtain bands after autoradiography. The filters were developed and analyzed as before. Arrows indicate the size in kilobases of the fragments. Fragments in panels B and D of 3.4 and 2.7 kb are from sarcoma component.

terminated previously (17). This sequence, which represents the 5' terminal 144 bases of the GaLVH RNA genome, contains a KpnI recognition site 32 bases from the end of the RNA. Since the strong-stop cDNAs of SSAV and the different GaLV isolates are closely related by oligonucleotide fingerprints (7), SSAV 5'-terminal sequences might possibly contain the same site.

KpnI digestion of unintegrated SSAV DNA gave four bands of 3.6, 2.7, 2.15, and 0.6 kb (Fig. 3, lane A). Identical bands were generated upon digestion of integrated SSAV proviral DNA (Fig. 3, lane B) except that the 0.6-kilobase pair band was absent, suggesting that it is attached to host DNA sequences. Therefore, one of the KpnI sites is 0.6 kb from one end of the viral DNA. Since the other three fragments are all internal and the other end fragment was not detected in KpnI digests of unintegrated SSAV provirus, it is likely that there is a KpnI site at or very near the other end of the proviral DNA. Since the sum of the four KpnI fragments was only 9.05 kb, and the full length provirus was 9.15 kb, there is probably a KpnI site within about 0.1 kb from the other end of the provirus. A 0.1-kb fragment would not be detectable in this hybridization system.

The internal KpnI fragments were ordered by partial and double digestions, as described above for BamHI and HindIII. A partial KpnI digest of unintegrated SSAV proviral DNA is shown in Fig. 3, lane C. The fragment sizes are 9.15, 6.7, 5.75, 4.9, 4.3, 3.6, 2.7, 2.15, and 0.6 kb. The complete digest fragment sizes are 3.6, 2.7, 2.15, and 0.6 kb. The largest bands in the
TABLE 1. Relative labeling of SSAV restriction fragments with 3' and total RNA

<table>
<thead>
<tr>
<th>Restriction nuclease</th>
<th>Fragment size (kb)</th>
<th>Relative intensity&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Relative intensity&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total RNA</td>
<td>3' probe RNA</td>
</tr>
<tr>
<td>BamHI</td>
<td>4.1</td>
<td>43.6</td>
<td>40.0</td>
</tr>
<tr>
<td></td>
<td>3.3</td>
<td>36.4</td>
<td>56.5</td>
</tr>
<tr>
<td></td>
<td>1.55</td>
<td>20.0</td>
<td>0.3</td>
</tr>
<tr>
<td>HindIII</td>
<td>4.25</td>
<td>45.9</td>
<td>55.7</td>
</tr>
<tr>
<td></td>
<td>1.9</td>
<td>18.1</td>
<td>4.3</td>
</tr>
<tr>
<td></td>
<td>1.6</td>
<td>23.2</td>
<td>35.7</td>
</tr>
<tr>
<td></td>
<td>1.4</td>
<td>12.7</td>
<td>4.3</td>
</tr>
</tbody>
</table>

<sup>a</sup> Size of the fragments as indicated by the arrows in Fig. 2.

<sup>b</sup> Areas under the densitometric peaks were determined by weight measurement after the peaks were traced out and cut with scissors. The weights were added and normalized for each tracing so that the total intensity was 100. The numbers indicate the relative percentage of each peak by weight.

partial digest are undigested linear SSAV (9.15-kb) and SSV (6.6-kb) proviral DNA. The presence of a 5.75-kb fragment indicated that the 3.6- and 2.15-kb fragments are adjacent. The presence of a 4.9-kb fragment indicated that the 2.70-kb fragment is on the other side of the 2.15-kb piece. This interpretation is strengthened by the lack of a 6.3-kb piece, which would be present if the two largest fragments were adjacent. The 0.6-kb end fragment is adjacent to the 3.6-kb piece, as indicated by the presence of a 4.3-kb fragment in the partial digest.

Double digests with KpnI and BamHI confirmed this assignment and indicated that the KpnI restriction map is 5'-0.6, 3.6, 2.15, 2.70-3'. Since the 3.6-kb KpnI fragment was not cleaved by BamHI (not shown). The position of the 3.6-kb fragment next to the 0.6-kb 5' end piece is further confirmed by its hybridization to strong-stop cDNA, discussed below.

The occurrence of KpnI sites close to both ends of the SSAV proviral DNA suggested that these sites are within the LTRs of the proviral DNA. If so, the size of the LTR would be a sum of the two end fragments, i.e., about 0.7 kb. This is similar to the LTR sizes which have been reported for feline leukemia virus (27) and Moloney-murine leukemia virus (37). A more precise placement of the KpnI site in the LTR was obtained by comparing the fragments generated by two enzymes which do not cut near the ends of the provirus (XhoI and Sall) with those generated by KpnI, by using 5' cDNA (see above for preparation) as a probe. The 5' cDNA hybridized strongly to both end fragments of XhoI and Sall digests (Fig. 4; Table 2), confirming that the SSAV provirus is terminally redundant and indicating that the terminal redundancy contains sequences derived from both the 3' (see results with the 3' probe and the BamHI and HindIII digests) and the 5' ends of the viral RNA. In contrast, only the left-hand (3.6 kb), internal (KpnI) fragment was strongly labeled with 5' cDNA. This suggests that KpnI cuts near the junction of the 3'- and 5'-derived sequences within the LTR, and that the 0.6-kb left-hand end fragment and the 2.7-kb right-hand internal fragment contain predominantly 3'-derived sequences. Indeed, the 3'-specific probe labeled only the latter fragment in digests of integrated proviral DNA (Fig. 5). The right-hand end fragment, consisting of approximately 0.1 kb of 5'-derived sequences, was not detectable with any cDNA. A rehybridization of these filters to total cDNA showed that the selective labeling with 5' and 3' probes was not due to a variation in the efficiency of transfer between different fragments (Fig. 4 and 5; Table 2). The structure of the SSAV linear unintegrated and integrated proviruses is therefore 3',5'-unique-3',5', similar to findings with other type-C viruses (6, 37). The proposed proviral structure is shown in Fig. 6.

The location of these KpnI sites in the SSAV LTR closely corresponds to that predicted from strong-stop sequence data with GalV<sub>H</sub> (17). Since strong-stop sequences have been shown to be highly conserved among this group of viruses

FIG. 3. KpnI digests of SSAV unintegrated and integrated DNA. Lane A, Complete KpnI digest of an unintegrated SSAV-SSV provirus. Fragment sizes are 3.6, 2.7, 2.15 and 0.6 kb. Lane B, KpnI digest of an SSAV-SSV integrated provirus. The fragment sizes are identical to those in lane A except that the 0.6-kb fragment is absent. Lane C, Partial KpnI digest of an SSAV-SSV integrated proviral DNA. Five fragments (9.15, 6.6, 5.75, 4.9, and 4.3 kb) are present which are not found in lane A.
(7) the KpnI LTR sites might possibly be conserved among all the members of the woolly monkey-gibbon group of retroviruses. Using experiments and logic similar to that used for SSAV, we found proviral DNA from all four GaLV isolates to have Kpn restriction sites within the 5'-derived sequences of the LTRs.

Based on the logic used above (namely, that internal fragments from an integrated provirus add up to the size of an unintegrated provirus minus 1 LTR or 8.4 kb), three other enzymes (SmaI, SstI, and PstI) also cut within the LTR of SSAV. The SmaI LTR sites map very near the KpnI LTR sites (Fig. 5), and as with KpnI, the left-most and right-most internal SmaI fragments are labeled with the 5'-(not shown) and 3'-specific (Fig. 5) probe, respectively. An SmaI site is also present in the GALVH strong-stop sequence (17) and is located only a few bases from the KpnI site, providing further evidence of the conservation of LTR sequences derived from the 5' portion of the RNA of these viruses. The SstI LTR site appears to be 80 to 100 bases to the left of the KpnI site and just outside the 5'-derived region. A 3' probe labeled both internal SstI fragments (Fig. 5). The PstI site is about 500 bases to the left of the KpnI site, near the junction of the right-hand LTR and the unique region of the provirus. A 3' probe specific for the first 200 to 250 bases of the RNA labeled only the left-hand internal PstI fragment (Fig. 5), indicating that the PstI LTR site is well into the 3'-derived sequences. The maps for the sites of these enzymes are shown in Fig. 6 and 7. Similar experiments with GaLVBr and GaLVSEATO indicated that SmaI, SstI, and PstI cut at the same sites within their proviral LTR sequences (Fig. 6 and Fig. 7, panels B and C). LTR sites for these enzymes and KpnI were also found for two other GaLV isolates, GaLVH (3) and GaLVSe (13) (not shown).

Maps of SSAV provirus with other restriction nucleases. SSAV proviral DNA was cleaved once with EcoRI, and fragments of 5.4 and 3.75 kb were produced. The 2.15-kb KpnI fragment from integrated provirus was cleaved by EcoRI. Since the KpnI map is 5'-0.6, 3.6, 2.15, 2.7, 0.1-3', the EcoRI site must be 5.4 kb from the 5' end.

**FIG. 4.** Hybridization of digestion fragments of Sall, XhoI, and KpnI to 5' and total cDNA. Unintegrated SSAV-SSV proviral DNA was digested with

\[ S\text{ali} \text{(panel A)}, \ Xh\text{oi} \text{(panel B)}, \text{or} \ KpnI \text{(panel C)}, \text{electrophoresed on agarose, transferred to nitrocellulose, and hybridized to} 5' \text{[}^{32}\text{P}]\text{cDNA as described in the text. After the bands were visualized by autoradiography and quantitated by scanning densitometry, the blots were washed and then rehybridized with total} \text{[}^{32}\text{P}]\text{cDNA as described for Fig. 2.} \] 

5' cDNA bands are represented by dashed lines, and bands from total cDNA are represented by a solid line. The numbers by the arrows indicate the size of the fragments in kilobases.
TABLE 2. Relative labeling of SSAV restriction fragments with 5' and total cDNA

<table>
<thead>
<tr>
<th>Restriction nuclease</th>
<th>Fragment size (kb)</th>
<th>Relative intensitya</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total cDNA</td>
<td>5' cDNA</td>
</tr>
<tr>
<td>Sall</td>
<td>5.4</td>
<td>65.5</td>
</tr>
<tr>
<td></td>
<td>2.7</td>
<td>17.8</td>
</tr>
<tr>
<td></td>
<td>1.1</td>
<td>16.7</td>
</tr>
<tr>
<td>Xhol</td>
<td>5.2</td>
<td>49.7</td>
</tr>
<tr>
<td></td>
<td>1.7</td>
<td>22.7</td>
</tr>
<tr>
<td></td>
<td>1.4</td>
<td>17.3</td>
</tr>
<tr>
<td></td>
<td>0.95</td>
<td>10.3</td>
</tr>
<tr>
<td>KpnI</td>
<td>3.6</td>
<td>49.7</td>
</tr>
<tr>
<td></td>
<td>2.7</td>
<td>16.7</td>
</tr>
<tr>
<td></td>
<td>2.15</td>
<td>28.2</td>
</tr>
<tr>
<td></td>
<td>0.6</td>
<td>5.4</td>
</tr>
</tbody>
</table>

a Size of the fragments as indicated by the arrows in Fig. 4.
b Areas under the densitometric peaks were quantitated and normalized as described for Table 1. The numbers indicate the relative intensity of each peak.

Sall cleaved unintegrated SSAV proviral DNA into three fragments, 5.4, 2.65, and 1.15 kb. Only the 1.15-kb band was found in Sall digests of integrated proviral DNA. Therefore, the 1.15-kb Sall fragment is internal. Digestion of integrated SSAV proviral DNA with KpnI and Sall together yielded fragments of 3.6, 2.70, and 1.1 kb. The 2.15-kb KpnI fragment was cleaved, indicating that a Sall site occurs in the center of this fragment, generating a doublet at 1.1 kb. Thus, the 5.4-kb fragment corresponds to the 5' end of the viral RNA.

Analogous methods were used for four more enzymes (BglII, Xhol, XbaI, SstII). Detailed descriptions of the data are omitted. The maps for all 12 restriction enzymes used are shown in Fig. 7 (panel A).

Maps of GaLV provirus with other restriction nucleases. Using similar experimental approaches, we constructed proviral maps of two of the isolates of GaLV (SEATO and Br) with eight enzymes (KpnI, PstI, SstI, Sall, Xhol, HindIII, BamHI, and EcoRI) to compare them with that of SSAV. The identical KpnI digestion patterns of the unintegrated and integrated proviruses of SSAV and these two GaLV isolates allowed us to orient the maps obtained with GaLV to the viral RNA. This orientation was confirmed by the use of 3' [32P]cDNA probes as described in the legend to Fig. 5. These maps are presented in Fig. 7 (panels B and C) and show that there is only about 50% conservation of non-LTR restriction nuclease sites between any two of these viruses and that only a few sites are present for all three. The low degree of conservation of restriction sites within the unique portion of these proviruses stands in marked contrast to sites within the LTR region and suggests that the unique sequences are generally much more variable than those of the LTRs.

FIG. 5. Hybridization of digestion fragments of KpnI, Smal, SstI, and PstI to 3' and total cDNA. Unintegrated linear proviral SSAV DNA was digested with the indicated enzymes, electrophoresed on agarose, transferred to nitrocellulose, and hybridized with 3' [32P]cDNA which was prepared by transcribing SSAV 3' RNA with avian myeloblastosis virus reverse transcriptase, by using oligo(deoxyribosylthymine)6CG as a primer and selecting cDNA 200 to 250 bases long on an 8% polyacrylamide slab gel as described in the text for strong-stop cDNA (panel A). After autoradiography, the same filter was then hybridized with total [32P]cDNA, and the bands were visualized by autoradiography (panel B).

Sizes of the fragments are as indicated in kilobases.
DISCUSSION

Restriction endonuclease maps of integrated and unintegrated proviral DNA from SSAV, GaLV_{SEATO}, and GaLV_{Br} were prepared and oriented 5' to 3' with respect to the viral genomic RNA, by using a combination of partial digests, double digests, and hybridization with end-specific probes. Several features of the resultant maps are of interest.

All of the above-mentioned proviruses are terminally redundant. Sequences derived from both the 3' and 5' ends of the RNA are present at both ends of the linear unintegrated provirus of SSAV. The LTRs of the provirus are approximately 700 base pairs long and consist of about 550 base pairs which derive from the 3' end of the viral RNA and about 150 base pairs which derive from the 5' end of the viral RNA. The orientation of the proviral sequences is 3',5'-unique-3',5'. Similar structures have been reported for linear DNA genomes of both murine (6, 37) and avian (10, 11, 26) retroviruses. Here, we have shown that all of the restriction endonuclease sites found within the LTR regions of

![Image of restriction endonuclease maps]

**FIG. 6.** Structural representation of the SSAV provirus. Proviral DNA is oriented above with respect to the homologous sequences in the RNA. The solid line represents the unique internal portion of the provirus, and the boxes represent the LTRs. The solid portion of the block represents sequences derived from the 5' 140 to 145 bases of the RNA (U5 and R) or strong-stop sequence, and the crosshatched portion of the block represents the U3 sequences deriving from the 3' region of the RNA other than the R sequences. Restriction sites are shown on the expanded scale box in the lower part of the figure. These sites are the same for all GaLV isolates tested. The polyadenylic acid [poly(A)] tract is not drawn to scale.

**FIG. 7.** Restriction endonuclease maps of linear proviral DNA of SSAV, GaLV_{SEATO}, and GaLV_{Br}. Boxes at the proviral termini indicate the extent of the LTR sequences derived from 3' and 5' regions of viral RNA. Calibration is in kilobases 5' to 3' with respect to viral RNA. Only _SmaI_ sites within the LTRs have been mapped for the two GaLV isolates. GaLV_{SEATO} _EcoRI_ sites have not been mapped. GaLV_{SEATO} has two species of both integrated and unintegrated proviruses, the minor of which contains an _SalI_ site at 5.3 map units. Abbreviations: _K_, _KpnI_; _S_, _SmaI_; _Si_, _SstI_; _P_, _PstI_; _Xb_, _XbaI_; _Xh_, _Xhol_; _Bg_, _BglII_; _B_, _BamHI_; _H_, _HindIII_; _E_, _EcoRI_; _SII_, _SstII_; _Sa_, _SalI_.

Although much of the information required for the complete understanding of the colinear relationship between viral and cellular DNA sequences has been obtained, several gaps remain. These gaps include a lack of information regarding the relative extents of integration as well as the degree and nature of the alterations that have occurred in the LTRs. The availability of the complete DNA sequence of the type C retrovirus MoMuLV provides a valuable model system for approach ing these problems. The MoMuLV genome is approximately 8,500 nucleotides and has been subcloned into a number of different plasmid vectors. The sequence of MoMuLV, including the LTRs, has been determined by using these subclones as probes in hybridization experiments. This sequence information is being used to define the structure and function of the LTRs and to analyze the relationship between the LTRs and the viral genome.
these viruses are completely conserved, confirming and extending previous results on the conservation of 5' sequences (7) to include the 3' sequences. It appears then that the LTR sequences play an important functional role in the life cycle of these viruses. It is interesting that murine type C viruses, which are believed to have a common progenitor with the infectious primate type-C viruses, also contain the same four restriction sites within their LTRs (20, 29, 31).

The LTRs are separated by about 7.75 kilobase pairs of sequences unique to the internal portion of the viral RNA. The unique part of the provirus consists of sequences in the right-hand 30% which contains the most variable region among the genomes of these viruses (22). The left-hand 70% appears very closely related in liquid hybridization experiments, but contains widely distributed small differences, as shown by restriction endonuclease mapping.

These maps will be useful for several kinds of future experiments on these viruses. For instance, they guided us in the strategy of cloning the SSV-SSAV genomes (5) and will help in cloning the genomes of the GaLV isolates. They should also help in the comparative analyses of other isolates of GaLV and allow a detailed characterization of partial proviruses which occur under natural conditions in gibbon apes (23, 36) or of proviral sequences in tissue from infected gibbon apes when a virus cannot be isolated or is defective. Finally, they may aid in the characterization of leukemia or disease-specific sequences which have been reported in some GaLV isolates (30).

ADDITIONAL IN PROOF
Recent experiments indicate that at least some preparations of GaLVSEATO proviral DNA lack an SsrI site within the LTR.

LITERATURE CITED
19. Pennas, S. 1969. Preparation of purified nucleic and nucleo-

Downloaded from http://jvi.asm.org/ on November 6, 2017 by guest


