JOURNAL OF VIROLOGY, May 1984, p. 451–456
0022-538X/84/050451-06$02.00/0

Vol. 50, No. 2

Monkey B-Lymphotropic Papovavirus DNA: Nucleotide Sequence of the Region Around the Origin of Replication

AKEMI FURUNO, TATSUO MIYAMURA, AND KUNITO YOSHIIKE*

Department of Enteroviruses, National Institute of Health, Kamiosaki, Shinagawa-ku, Tokyo 141, Japan

Received 15 September 1983/Accepted 10 January 1984

We have determined the nucleotide sequence of the *Hin*dIII-B DNA segment of African green monkey B-lymphotropic papovavirus (LPV), which shows a highly restricted host range and whose genome is a 5.1-kilobase-long circular DNA. The segment, consisting of 1,123 base pairs, contained the origin of DNA replication, the putative control region for early transcription, and the region probably coding for the amino-terminal portion of T antigens. The symmetrical region at the center of replication origin, 5'-GAGGC CA GGGGCCCC TA GCCTC-3' (on the L strand), has diverged in the central portion from the corresponding regions of primate polyomaviruses simian virus 40, BK virus, and JC virus, but resembles that of mouse polyomavirus. The structure of the control region upstream of the replication origin was unique to LPV and contained several repeated sequences, the longest of which were two 60-base-pair tandem repeats. The amino-terminal region common to LPV small T and large T antigens showed some homology (41%) in the deduced amino acid sequence to that of both simian virus 40 and the mouse polyomavirus. Like other polyomaviruses, the probable carboxyl-terminal region unique to LPV small T antigen contained two sets of the Cys-x-Cys-x-x-Cys structure. These data show that, despite the unique structures in the control region, LPV is evolutionally related to the mouse polyomavirus and to simian virus 40.

Lymphotropic papovavirus (LPV), isolated from a B-lymphoblastoid cell line of the African green monkey (41), shows a highly restricted host range. It grows only in certain monkey or human B-lymphoblastoid cell lines, but not in most of the T and null lymphoblastoid cell lines (3, 4, 33). Although serological surveys have shown that the yet unidentified viruses antigenically related to LPV must occur widely among humans, apes, and monkeys (3, 33), the pathogenicity of LPV is still unknown. For its special host range and prevalence of antibody to it in primates, LPV is distinct from other primate polyomaviruses.

Nondefective monkey LPV DNA has been characterized by molecular cloning from heterogeneous DNA grown in human B-lymphoblastoid BJA-B cells (33). Circular 5.1-kilobase (kb)-long LPV DNA begins to replicate bidirectionally from a point located near its unique *Bam*HI cleavage site (16). Hybridization experiments with specific DNA fragments revealed partial homology over the entire genomes between LPV DNA and simian virus 40 (SV40) or BK virus (BKV) DNA, indicating that LPV is evolutionally related to these polyomaviruses, and the LPV genome can be aligned to those of SV40 and BKV with the replication origin as a reference point (16). It is now possible to deduce the correlation between the physical and functional maps of the LPV genome, if LPV has the genomic organization common to the polyomaviruses.

The noncoding region around the replication origin, whose nucleotide sequences are the least homologous even among closely related primate polyomaviruses SV40, BKV, and JC virus (JCV) (9, 24, 27, 35, 39), contains the control signals affecting diverse biological properties of polyomaviruses including the host range of mouse polyomavirus (10, 18, 28), SV40 (5, 19), and JCV (40) and the transforming capacity of BKV (37). Detailed analyses of the SV40 genome have shown that the three 21-base-pair (bp) repeats upstream from the replication origin function as the important elements of the promoter (5), and that the two 72-bp tandem repeats further upstream have the enhancing and host-discriminatory functions for the early viral transcription in vivo (1, 5, 13, 38). Since LPV has a special host range among polyomaviruses, the anatomy of the structure of the region encompassing the origin of DNA replication is interesting and essential for further analyses of possible unique signals of LPV controlling expression of the viral genome in B-lymphoblastoid cells.

In the present study we determined the nucleotide sequence of *Hin*dIII-B fragment of LPV DNA. Within this segment of 1,123 bp, we located the origin of DNA replication, the putative control region for early transcription, and the probable coding region for the early viral proteins.

MATERIALS AND METHODS

LPV DNA. A recombinant plasmid, pL02, which has a nondefective 5.1-kb LPV DNA insert at the *Bam*HI site of plasmid pBR322 (32), and its derivatives, pL026 and pL027, were used. Recombinants pL026 and pL027 contain *Bam*HI-*Hin*dIII-B_2 and -B_1, respectively, of LPV DNA between *Bam*HI and *Hin*dIII sites of pBR322. Molecular cloning, preparation of recombinant DNA, and the cleavage maps of LPV DNA were described previously (33).

Enzymes. Restriction endonucleases, purchased from Bethesda Research Laboratories, Rockville, Md. (*Alu*I, *Ava*I, *Ava*II, *Bam*HI, *Dde*I, *Eco*RI, *Hin*dIII, *Hpa*II, and *Sau*96I), from New England BioLabs, Beverly, Mass. (*Bst*NI, *Hin*fI, *Nco*I, and *Rsa*I), and from Takara Shuzo Co., Ltd., Kyoto, Japan (*Hae*II, *Hae*III, and *Hha*I), were used as recommended by the manufacturers. Bacterial alkaline phosphatase and T4 polynucleotide kinase were purchased from Worthington Biochemical Corp., Freehold, N.J., and Boehringer Mannheim GmbH Biochemica, Mannheim, West Germany, respectively.

Approximate location of origin of LPV DNA replication. LPV DNA, either isolated from the recombinant plasmid pL02 or extracted from LPV-infected BJA-B cells (33), was

* Corresponding author.

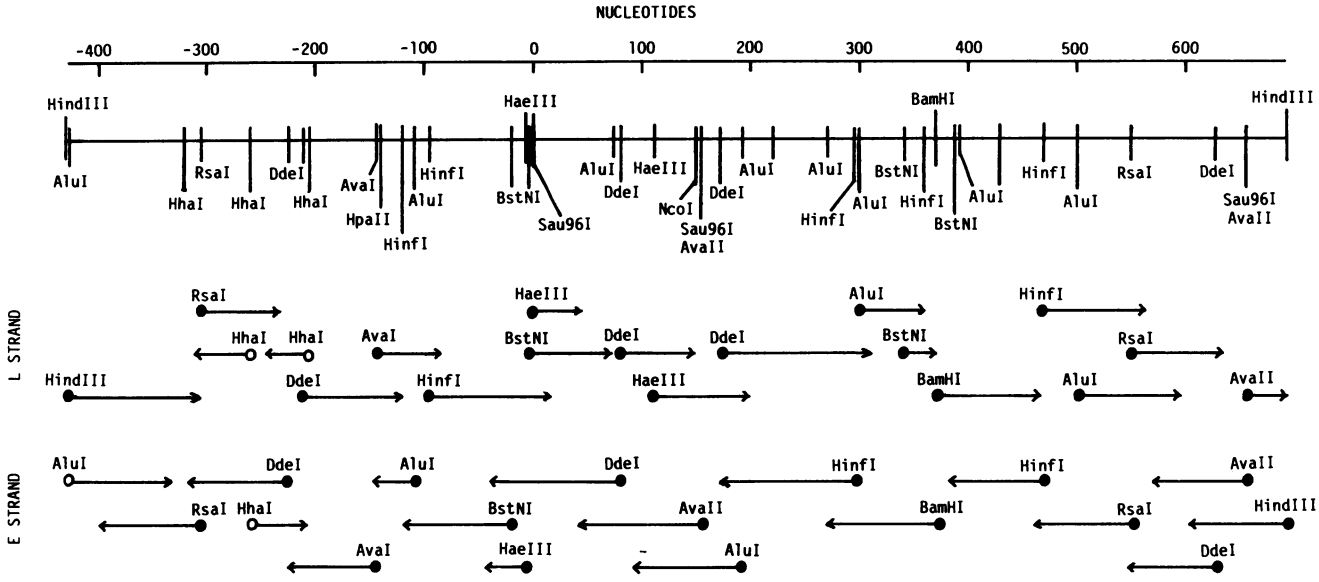

FIG. 1. Schematic representation of the strategy used to sequence the *Hin*dIII-B segment of LPV DNA. *Hin*dIII cleaves LPV DNA at two sites, generating 4-kb (*Hin*dIII-A) and 1.1-kb (*Hin*dIII-B) fragments (33). The relative positions of restriction endonuclease sites were determined either by comparative sizing of digested fragments or directly by sequencing. Fragments were labeled at 5′ ends (●) or 3′ ends (○). The arrows indicate the direction and approximate extent of reading. *Hae*II sites not shown here overlap *Hha*I sites. The DNA sequencing shown in Fig. 2 revealed an *Ava*II cleavage site (5′-GGTCC-3′) at nucleotides 334 to 338, partially overlapping a *Bst*NI site (5′-CCAGG-3′) at nucleotides 337 to 341. This *Ava*II sequence was insensitive to *Ava*II and *Sau*96I, probably because of methylation occurring at the second C in a stretch of 5′-CCAGG-3′. pBR322 DNA also has an *Ava*II site partially overlapping a *Bst*NI site in a stretch of 5′-GGTCCTGG-3′ at nucleotides 1,438 to 1,445 (32). In accordance with Sutcliffe (32), *Ava*II did not cut our pBR322 DNA at this *Ava*II site.

cleaved with *Bam*HI, *Hin*dIII, and *Hpa*II, each alone or with appropriate combination, and subjected to electrophoresis in 1.0% agarose as described previously (33). Separated DNA fragments in the gels were transferred to membrane filters by the method of Southern (31) and hybridized to purified BKV DNA *Hin*dIII-C fragments, which consist of 555 bp and contain the BKV DNA replication origin and its surrounding area (27, 39), labeled with [α-^{32}P]dCTP (specific activity, 600 Ci/mmol; Amersham International plc, Buckinghamshire, England) by nick translation (25). The method of hybridization was as described by Howley et al. (15), with various concentrations (20, 40, and 60%) of formamide used to change the stringency of hybridization conditions.

DNA sequencing. Nucleotide sequences were determined by the chemical method of Maxam and Gilbert (21). LPV DNA digested with appropriate restriction enzymes was subjected to preparative electrophoresis in 5% polyacrylamide–0.25% bisacrylamide gels. When necessary, double-stranded fragments isolated from gels were cleaved with the second restriction enzymes. Specific DNA fragments were labeled at their 5′ ends with [γ-^{32}P]ATP (specific activity, 3,000 or 6,000 Ci/mmol; Amersham International) and T4 polynucleotide kinase or labeled at their 3′ ends with [α-^{32}P]ddATP (specific activity, 3,000 Ci/mmol; Amersham International) with a 3′ end-labeling kit (Amersham International). Fragments labeled at one end were prepared by strand separation (in 8% polyacrylamide–0.16% bisacrylamide) or cleavage with an appropriate restriction enzyme and were submitted to the four sets of partial degradation reaction of A>C, G, C, and C+T. The degradation products were run on 0.5-mm-thick 12% polyacrylamide–0.6% bisacrylamide–8 M urea gels. Autoradiography was performed at −80°C with an intensifying screen.

RESULTS

For determination of the segment containing the origin of LPV DNA replication, specific LPV DNA fragments immobilized on membrane filters were hybridized to 0.56-kb ^{32}P-labeled BKV DNA fragments (specific activity, 1.3×10^8 cpm/μg of DNA) containing the origin of BKV DNA replication. Homology was detected only under the lowest stringency conditions (20% formamide, T_m − 50°C) between BKV DNA fragments and LPV DNA *Bam*HI-*Hin*dIII-B_1 or LPV DNA *Bam*HI-*Hpa*II-B_2 (data not shown), indicating that the LPV DNA replication origin was probably located between 0 and 0.1 map unit from the unique *Bam*HI cleavage site. Therefore, the *Hin*dIII-B segment (1.1 kb long) of LPV DNA containing *Bam*HI site was chosen for DNA sequencing.

Figure 1 illustrates the cleavage sites of various restriction endonucleases in the LPV DNA *Hin*dIII-B segment and schematic representation of the strategy for sequencing, and Fig. 2 summarizes the results obtained. The LPV sequence was compared with those around the DNA replication origin in SV40 (24, 35, 36), BKV (27, 39), and JCV (9, 22), and the mouse polyomavirus (7, 8, 17, 29, 30).

Origin of DNA replication. We have located the origin of LPV DNA replication between nucleotides −32 and 49 (Fig. 2) from the structural similarity among polyomaviruses. The core of SV40 DNA replication origin, which has been extensively studied (2, 14), consists of a 17-bp AT-rich sequence, a 27-bp perfect palindrome having a two-fold rotational symmetry, a 15-bp palindrome, and a 17-bp true palindrome. Similar structures were also found in the LPV DNA (Fig. 3).

In the replication origin, a 27-bp symmetrical structure is the area where the sequence homology is most highly conserved among SV40, BKV, and JCV DNAs and probably

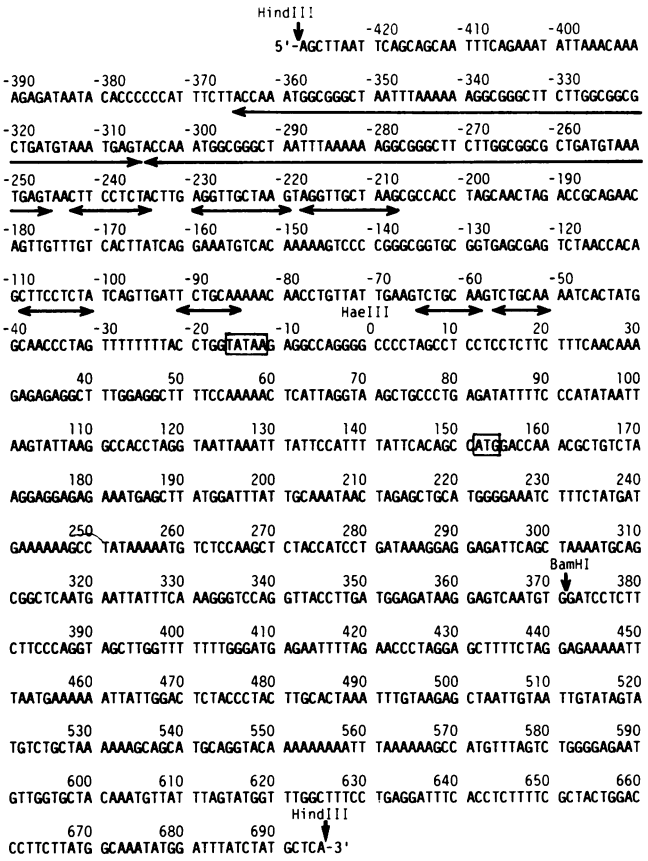

FIG. 2. Nucleotide sequence of LPV DNA, L strand of *Hin*dIII-B-segment. The sequence has the same polarity as the early mRNA in the 5' to 3' direction. One of the *Hae*III cleavage sites within replication origin has been arbitrarily taken as the zero point for numbering nucleotides. Some signals are boxed with a solid line, the TATA box at nucleotides −16 to −12 and the initiation codon for T antigen mRNA at nucleotides 152 to 154. The arrows indicate some of the repeated sequences.

represents the center of the replication origin. Whereas the highly conserved sequence of SV40 is 5'-GAGGC C GAGGC C GCCTC G GCCTC-3' on the L strand (24, 35), the corresponding sequence of LPV was 5'-GAGGC CA GGGGCCCC TA GCCTC-3' (nucleotides −11 to 11 in Fig. 2). This sequence of LPV resembled the recently reported counterpart of the mouse polyomavirus, 5'-GAGGC CG GGGGCCCC TG GCCTC-3' (17). Except for an identical stretch of 8 T's (nucleotides −30 to −23 in Fig. 2), other structures were loosely conserved in LPV DNA (Fig. 3).

The arrangement of possible large T antigen recognition pentanucleotides around the LPV DNA replication origin is shown in Fig. 3. The recognition pentanucleotides of the consensus family 5'-(G>T)(A>G)GGC-3' of SV40 (6) were found in LPV DNA distributed in the regions corresponding to T antigen binding sites I and II of SV40 (34). Between nucleotides −12 and 12 four sets of pentanucleotides were oriented as inverted repeats, and the two sets of 5'-GGGGC-3' with opposite polarity at the center were partially overlapping. This region apparently corresponds to site II of SV40. The region between nucleotides 34 and 49, including two sets of 5'-GAGGC-3', probably corresponds to site I of SV40. In the area topographically corresponding to site III of SV40, LPV had no sets of the consensus pentanucleotides.

Early transcriptional control region. Topographically, the replication origin and its upstream area in the *Hin*dIII-B segment corresponds to the early transcriptional control region of SV40 (5), which consists of two 72-bp tandem repeats and three 21-bp repeats followed by the core structure of the replication origin. Although the structure of LPV DNA in this region (nucleotides −428 to −1) was different from those of SV40 (35), BKV (27), JCV (9), and the mouse polyomavirus (30), it contained its characteristic sets of repeated sequences.

The longest of these are two 60-bp tandem repeats at nucleotides −365 to −306 and nucleotides −305 to −246 (Fig. 2). Within the 60-bp tandem repeats were GC-rich repeats; four sets of 5'-GGCGGGC-3' (nucleotides −358 to −352, nucleotides −339 to −333, nucleotides −298 to −292, and nucleotides −279 to −273) and two sets of 5'-GGCGGCGC-3' (nucleotides −327 to −320, and nucleotides −267 to −260). Other sets of two repeats were 5'-CTTCCTCTA-3' (at nucleotides −243 to −235 and nucleotides −109 to −101), and 5'-AGGTTGCTAAG-3' (at nucleotides −230 to −220 and nucleotides −218 to −208). Near the replication origin there were three repeats of 5'-TCTGCAA-3' at nucleotides −91 to −85, nucleotides −65 to −59, and nucleotides −57 to −51.

Coding region for viral early proteins. Like other polymaviruses, LPV is believed to code for T antigens. Although the 84K protein found in the proteins synthesized in vitro (26) may be LPV large T antigen, the presence of small T antigen remains to be proved.

The probable coding region for T antigens initiates from the ATG at nucleotides 152 to 154. In the region downstream from the replication origin, this ATG was the first initiation codon followed by the longest open reading frame within the *Hin*dIII-B segment. The deduced amino acid sequence, together with those of SV40 and the mouse polyomavirus, is shown in Fig. 4.

By analogy to other polyomaviruses, the amino-terminal region of LPV small T antigen is believed to be shared by large T antigen. From the similarity to consensus sequences for splicing (23), LPV DNA has two possible splice donor

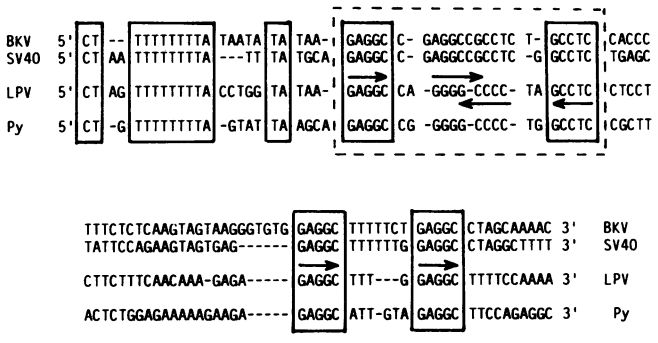

FIG. 3. Comparison of the sequences of the replication origins of LPV, SV40, BKV, and mouse polyomavirus DNAs. The sequence of SV40 includes a 17-bp AT-rich stretch, a 27-bp perfect palindrome, a 15-bp palindrome, and a 17-bp true palindrome. Conserved stretches are boxed with a solid line. The symmetrical region (included in the 27-bp palindrome in SV40 DNA) is boxed with a dashed line. The arrows indicate distribution of T-antigen recognition pentanucleotides (described previously for SV40 DNA [6, 34]) in the LPV DNA replication origin. In the LPV sequence, the first C and the last A are nucleotides −34 and 58, respectively, in Fig. 2. BKV and SV40 sequences are from published data (24, 27, 35). The mouse polyomavirus (Py) sequence is from Katinka and Yaniv (17).

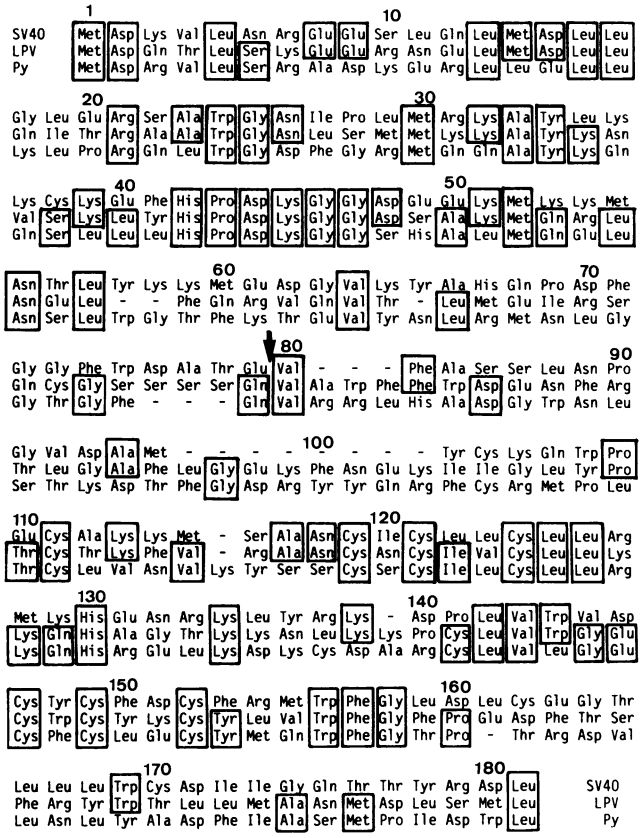

FIG. 4. Amino acid sequences of the amino-terminal portions of SV40, LPV, and mouse polyomavirus T antigens. The sequence of LPV was deduced from the nucleotide sequence from nucleotides 152 to 694 in Fig. 2. The sequences for SV40 and mouse polyomavirus (Py) are from published data (8, 29, 36). The greater parts of SV40 and mouse polyomavirus small T antigens consisting of 174 and 195 amino acids, respectively, are shown here. Gaps were left where necessary to maximize homology. The alignment between SV40 and mouse polyomavirus was modified from Seif et al. (27). The boxes indicate the homologous amino acids between LPV and SV40 or mouse polyomavirus. Numbers refer to amino acid positions for LPV. The first Met and the last Leu of LPV are numbered as positions 1 and 181, respectively. The arrow between Gln and Val (positions 79 and 80) of LPV divides the amino-terminal region common to the small T and large T antigens from the carboxyl-terminal region unique to the small T antigen.

sites for large T antigen mRNA between nucleotides 340 and 341 and between nucleotides 388 and 389 (Fig. 2).

DISCUSSION

We have located the origin of LPV DNA replication in the *Hin*dIII-B segment from the structural characteristics common to the replication origins of SV40, BKV, and mouse polyomavirus DNAs (Fig. 3). By analogy to these other viruses, identification of the replication origin in the LPV DNA sequence immediately predicts the orientation of the early and late regions of the LPV genome. Downstream from the replication origin we located a region probably encoding the amino-terminal portion of the T antigens of LPV. The predicted genomic organization of LPV is consistent with the deduced correlation of physical and functional maps to those of SV40 or BKV (16).

The structure of replication origin of monkey LPV DNA was similar to those of primate polyomaviruses SV40, BKV, and JCV. Probably the most important difference is in the region corresponding to the symmetrical 27-bp palindrome of SV40, which is most highly conserved in the cores of the SV40, BKV, and JCV origins. The highly conserved sequence of LPV DNA replication origin was most related to the mouse polyomavirus (Fig. 3).

The core region of replication origin contains a number of polyomavirus promoter elements. Topographically, TATAA at nucleotides −16 to −12 would appear to function as a Goldberg-Hogness box and position precisely (20) the 5′ ends of early mRNA. SV40 early mRNA starts preferentially at 22 bp downstream from its TATA box and reads GCCTCTGAGCTATTCCA (12). The corresponding sequence of LPV beginning 18 bp downstream from TATAA (nucleotides −16 to −12) reads GCCTCCTCCTCTTCTTT. Whether this is actually the 5′ end of LPV mRNA and the possible function of the second sequence (TATAAT) at nucleotides 94 to 99 (Fig. 2) remain to be determined.

The region upstream of the replication origin contains the signals controlling the early transcription, viability, host range, and transforming capacity of polyomaviruses (10, 11, 13, 18, 19, 28, 37, 38, 40). The structure of this region is different from virus to virus (9, 17, 24, 27, 30, 35, 39). The control region of SV40 consists of three 21-bp repeats, which are located just upstream of the replication origin and act as the major element of the early promoter, and two 72-bp tandem repeats, which are located upstream of the 21-bp repeats and function as the host-discriminating enhancer for early transcription (5, 13, 38). Although the comparable region of LPV (nucleotides −428 to −1 in Fig. 2) did not contain the same structure as SV40 has, it contained several sets of repeated sequences unique to LPV, the longest of which are 60-bp tandem repeats superficially resembling 72-bp repeats of SV40. The domains for promoter components and enhancer elements of LPV are yet to be determined.

Despite the superficial similarity, the nucleotide sequence of the LPV 60-bp element was different from that of the SV40 72-bp element, and the structures between the tandem repeats and the replication origin were different between LPV and SV40. The major components of SV40 early promoter consisting of three 21-bp repeats contain six GC-rich hexanucleotides (5′-CCGCCC-3′ on the L strand). The topographically comparable region of LPV had no 21-bp repeats and no GC-rich hexanucleotide sets. The only GC-rich stretch in this region was 5′-GTCCCCGGGCGGTGCGG-3′ at nucleotides −145 to −129 (Fig. 2). Instead, the 60-bp elements contained short characteristic GC-rich segments of 5′-GGCGGGC-3′ (on the L strand), which resembled, on the complementary strand, 5′-CCGCCC-3′ found in the major component of the SV40 promoter. Furthermore, the LPV 60-bp repeats did not contain a complete set of 5′-TGGA(T)A(T)A(T)-3′, which is believed to be the potential core sequence in the polyomavirus enhancers first found in the 72-bp enhancer element of SV40 (38). The LPV sequence resembling this core sequence, 5′-GGTGTGGAAAG-3′ (on the L strand), of the SV40 enhancer was located at nucleotides −119 to −109 (5′-CTAACCACAGC-3′ on the L strand in Fig. 2) with the opposite orientation, reading on the E strand, 5′-GCTGTGGTTAG-3′. Like the SV40 counterpart in the 72-bp element, this 5′-TGGTTA-3′ sequence of LPV (nucleotides −118 to −113 read on the E strand) was followed by a GC-rich stretch (nucleotides −145 to −121). In addition, the potential core sequence (38) of 5′-TGGA(T)A(T)A(T)-3′ was located at three more sites, in the noncoding region around LPV DNA replication origin, nucleotides −367 to −362,

nucleotides 50 to 55, and nucleotides 132 to 137, all with opposite orientation to those of SV40. One of the 60-bp repeats contained only a part of this core sequence (nucleotides −367 to −362 in Fig. 2).

Interestingly, the putative control region of LPV DNA contained some sequences resembling characteristic SV40 signals such as 5′-TGGAAA-3′ (for the enhancer) and 5′-CCGCCC-3′ (for the promoter), with the opposite orientation or with the exchanged nucleotides at different topographical sites. Furthermore, the LPV control region contained unique repeated sequences. Whether the SV40-like signals of LPV function like the SV40 counterparts in LPV-infected cells and whether unique repeats of LPV are related to any of the biological characteristics of LPV will be analyzed in the future studies.

If LPV had a small T antigen with a molecular weight similar to those of other polyomaviruses, the amino acid sequence shown in Fig. 4 would constitute the greater part of this protein. By analogy to SV40, the LPV large T antigen is expected to share the amino acid sequence of the amino-terminal half of small T antigen. The surrounding sequences of two possible splice donor sites for LPV large T antigen mRNA read GTCCAG/GTTACCT (nucleotides 335 to 347 in Fig. 2) and TCCCAG/GTAGCTT (nucleotides 383 to 395 in Fig. 2), both of which contain a stretch similar to the consensus sequence C(A)AG/GTA(G)AGT (23). Based on the topographical similarity among polyomaviruses (Fig. 4) and the similarity to the consensus sequence, we consider the site between nucleotides 388 to 389 (between Gln and Val at positions 79 and 80, respectively) to be the splice donor site for LPV large T antigen mRNA.

The homology of amino acid sequences between LPV and SV40 or mouse polyomavirus was higher in the amino-terminal region common to small T and large T antigens than in the carboxyl-terminal region unique to small T antigen (Fig. 4). The degree of homology in the common region was 41% (32 of 79 amino acids) both between monkey LPV and SV40 and between LPV and the mouse polyomavirus, whereas the degree of homology in the same region is more than 80% among primate polyomaviruses SV40, BKV, and JCV (22).

Like other polyomaviruses (7, 27), the unique region of LPV small T antigen contained two sets of the Cys-x-Cys-x-x-Cys structure (positions 119 to 124, and 147 to 152 in Fig. 4). These conserved stretches of amino acids may be related to the biological functions unique to small T antigens of polyomaviruses.

The present analyses of LPV DNA sequence have revealed that monkey LPV equally resembles both monkey and mouse polyomaviruses. LPV and the mouse polyomavirus share a common structure at the central, symmetrical portion of the DNA replication origin. The degree of homology in amino acid sequence in the amino-terminal half of the small T antigen between LPV and SV40 was similar to that between LPV and the mouse polyomavirus. The results show that the homology among LPV, SV40, and the mouse polyomavirus does not reflect the phylogenetic relationships of their natural hosts, African green monkey, rhesus monkey, and mouse, respectively. Nevertheless these data, together with those of the previous study (16), suggest that these viruses must have evolved from a common polyomavirus ancestor and that LPV diverged earlier than did SV40, BKV, and JCV.

In summary we have determined the nucleotide sequence of LPV DNA *Hin*dIII-B segment (1,123 bp) containing the origin of DNA replication. The data should facilitate studies to determine the domains of promoter and enhancer for early transcription and to identify the signals related to the special host range of LPV and will provide basis for further biological and biochemical studies of LPV.

ACKNOWLEDGMENTS

This work was supported in part by grants from Ministry of Education, Science, and Culture, and from Science and Technology Agency, Japan.

LITERATURE CITED

1. **Benoist, C., and P. Chambon.** 1981. *In vivo* sequence requirements of the SV40 early promoter region. Nature (London) **290**:304–310.
2. **Bergsma, D. J., D. M. Olive, S. W. Hartzell, and K. N. Subramanian.** 1982. Territorial limits and functional anatomy of the simian virus 40 replication origin. Proc. Natl. Acad. Sci. U.S.A. **79**:381–385.
3. **Brade, L., N. Müller-Lantzsch, and H. zur Hausen.** 1980. B-Lymphotropic papovavirus and possibility of infections in humans. J. Med. Virol. **6**:301–308.
4. **Brade, L., W. Vogl, L. Gissman, and H. zur Hausen.** 1981. Propagation of B-lymphotropic papovavirus (LPV) in human B-lymphoma cells and characterization of its DNA. Virology **114**:228–235.
5. **Byrne, B. J., M. S. Davis, J. Yamaguchi, D. J. Bergsma, and K. N. Subramanian.** 1983. Definition of the simian virus 40 early promoter region and demonstration of a host range bias in the enhancement effect of the simian virus 40 72-base-pair repeat. Proc. Natl. Acad. Sci. U.S.A. **80**:721–725.
6. **DeLucia, A. L., B. A. Lewton, R. Tjian, and P. Tegtmeyer.** 1983. Topography of simian virus 40 A protein-DNA complexes: arrangement of pentanucleotide interaction sites at the origin of replication. J. Virol. **46**:143–150.
7. **Friedmann, T., R. F. Doolittle, and G. Walter.** 1978. Amino acid sequence homology between polyoma and SV40 tumour antigens deduced from nucleotide sequences. Nature (London) **274**:291–293.
8. **Friedmann, T., A. Esty, P. LaPorte, and P. Deininger.** 1979. The nucleotide sequence and genome organization of the polyoma early region: extensive nucleotide and amino acid homology with SV40. Cell **17**:715–724.
9. **Frisque, R. J.** 1983. Nucleotide sequence of the region encompassing the JC virus origin of DNA replication. J. Virol. **46**:170–176.
10. **Fujimura, F. K., P. L. Deininger, T. Friedmann, and E. Linney.** 1981. Mutation near the polyoma DNA replication origin permits productive infection of F9 embryonal carcinoma cells. Cell **23**:809–814.
11. **Gheysen, D., A. Van de Voorde, R. Contreras, J. Vanderheyden, F. Duerinck, and W. Fiers.** 1983. Simian virus 40 mutants carrying extensive deletions in the 72-base-pair repeat region. J. Virol. **47**:1–14.
12. **Ghosh, P. K., and P. Lebowitz.** 1981. Simian virus 40 early mRNA's contain multiple 5′ termini upstream and downstream from a Hogness-Goldberg sequence; a shift in 5′ termini during the lytic cycle is mediated by large T antigen. J. Virol. **40**:224–240.
13. **Gruss, P., R. Dhar, and G. Khoury.** 1981. Simian virus 40 tandem repeated sequences as an element of the early promoter. Proc. Natl. Acad. Sci. U.S.A. **78**:943–947.
14. **Hay, R. T., and M. L. DePamphilis.** 1982. Initiation of SV40 DNA replication in vivo: location and structure of 5′ ends of DNA synthesized in the *ori* region. Cell **28**:767–779.
15. **Howley, P. M., M. A. Israel, M.-F. Law, and M. A. Martin.** 1979. A rapid method for detecting and mapping homology between heterologous DNAs. Evaluation of polyomavirus genomes. J. Biol. Chem. **254**:4876–4883.
16. **Kanda, T., K. Yoshiike, and K. K. Takemoto.** 1983. Alignment of the genome of monkey B-lymphotropic papovavirus to the genomes of simian virus 40 and BK virus. J. Virol. **46**:333–336.

17. **Katinka, M., and M. Yaniv.** 1983. DNA replication origin of polyoma virus: early proximal boundary. J. Virol. **47**:244–248.
18. **Katinka, M., M. Yaniv, M. Vasseur, and D. Blangy.** 1980. Expression of polyoma early functions in mouse embryonal carcinoma cells depends on sequence rearrangements in the beginning of the late region. Cell **20**:393–399.
19. **Laimins, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **79**:6453–6457.
20. **Mathis, D. J., and P. Chambon.** 1981. The SV40 early region TATA box is required for accurate *in vitro* initiation of transcription. Nature (London) **290**:310–315.
21. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
22. **Miyamura, T., H. Jikuya, E. Soeda, and K. Yoshiike.** 1983. Genomic structure of human polyoma virus JC: nucleotide sequence of the region containing replication origin and small-T-antigen gene. J. Virol. **45**:73–79.
23. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
24. **Reddy, V. B., B. Thimmappaya, R. Dhar, K. N. Subramanian, B. S. Zain, J. Pan, P. K. Ghosh, M. L. Celma, and S. M. Weissman.** 1978. The genome of simian virus 40. Science **200**:494–502.
25. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling deoxyribonucleic acid to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **113**:237–251.
26. **Segawa, K., and K. K. Takemoto.** 1983. Identification of B-lymphotropic papovavirus-coded proteins. J. Virol. **45**:872–875.
27. **Seif, I., G. Khoury, and R. Dhar.** 1979. The genome of human papovavirus BKV. Cell **18**:963–977.
28. **Sekikawa, K., and A. J. Levine.** 1981. Isolation and characterization of polyoma host range mutants that replicate in nullipotential embryonal carcinoma cells. Proc. Natl. Acad. Sci. U.S.A. **78**:1100–1104.
29. **Soeda, E., J. R. Arrand, N. Smolar, and B. E. Griffin.** 1979. Sequence from early region of polyoma virus DNA containing viral replication origin and encoding small, middle and (part of) large T antigens. Cell **17**:357–370.
30. **Soeda, E., J. R. Arrand, N. Smolar, J. E. Walsh, and B. E. Griffin.** 1980. Coding potential and regulatory signals of the polyoma virus genome. Nature (London) **283**:445–453.
31. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
32. **Sutcliffe, J. G.** 1979. Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43**:77–90.
33. **Takemoto, K. K., A. Furuno, K. Kato, and K. Yoshiike.** 1982. Biological and biochemical studies of African green monkey lymphotropic papovavirus. J. Virol. **42**:502–509.
34. **Tjian, R.** 1979. Protein-DNA interactions at the origin of simian virus 40 DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:655–662.
35. **Van Heuverswyn, H., and W. Fiers.** 1979. Nucleotide sequence of the *Hin*d-C fragment of simian virus 40 DNA. Comparison of the 5′-untranslated region of wild-type virus and of some deletion mutants. Eur. J. Biochem. **100**:51–60.
36. **Volckaert, G., A. Van de Voorde, and W. Fiers.** 1978. Nucleotide sequence of the simian virus 40 small-t gene. Proc. Natl. Acad. Sci. U.S.A. **75**:2160–2164.
37. **Watanabe, S., and K. Yoshiike.** 1982. Change of DNA near the origin of replication enhances the transforming capacity of human papovavirus BK. J. Virol. **42**:978–985.
38. **Weiher, H., M. König, and P. Gruss.** 1983. Multiple point mutations affecting the simian virus 40 enhancer. Science **219**:626–631.
39. **Yang, R. C. A., and R. Wu.** 1979. BK virus DNA: complete nucleotide sequence of a human tumor virus. Science **206**:456–462.
40. **Yoshiike, K., T. Miyamura, H. W. Chan, and K. K. Takemoto.** 1982. Two defective DNAs of human polyomavirus JC adapted to growth in human embryonic kidney cells. J. Virol. **42**:395–401.
41. **zur Hausen, H., and L. Gissmann.** 1979. Lymphotropic papovaviruses isolated from African green monkey and human cells. Med. Microbiol. Immunol. **167**:137–153.