Proc. Natl. Acad. Sci. USA
Vol. 78, No. 1, pp. 95–99, January 1981
Biochemistry

DNA sequences similar to those around the simian virus 40 origin of replication are present in the monkey genome

(simian virus 40 DNA/monkey DNA library/recombinant DNA/host–virus homology)

THOMAS F. MCCUTCHAN AND MAXINE F. SINGER

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Contributed by Maxine F. Singer, September 12, 1980

ABSTRACT **We report the molecular cloning of African green monkey genomic DNA segments that include regions of homology to the origin of replication of simian virus 40 (SV40). Three clearly different cloned segments 14 to 17 kilobase pairs (kb) long were isolated from a genomic library in λ phage. We estimate that each of the three is repeated fewer than four times in the monkey genome. The SV40-like regions represent a small portion of the cloned segments, and these regions cross hybridize only weakly with one another. One of the three segments is described here in detail. Although the entire segment occurs only once or twice in the monkey genome, it contains DNA sequences (other than the SV40-like sequences) that are repeated elsewhere in the genome including in the other two cloned segments. The homology to SV40 is contained within about 300 base pairs of monkey DNA and is limited to the region around the viral replication origin. The nucleotide sequence of the SV40-like region was determined. It contains a large number of short stretches homologous to three specific noncoding domains around the SV40 origin of replication: the 27-base-pair region of dyad symmetry, the first set of (short) repeats that occur just on the late side of the origin, and, further in the late direction, the two 72-base-pair-long repeats. Although these components are grouped in the monkey DNA, as they are in SV40 DNA, their relative juxtaposition is scrambled.**

The genomes of certain eukaryotic viruses and their respective uninfected hosts have been shown to carry homologous regions of nucleic acid sequence (1–3). Although sequence homology between the genomes of host and virus can be in either coding or noncoding regions or in both, genomic similarity thus far has been demonstrated only in coding regions (1, 2). Virus–host pairs carrying homologies within noncoding sequences (e.g., promoters for RNA transcription, origins of replication, termination sites for RNA polymerases) have not been documented. Common or similar control regions in noncoding sequences might be expected, however, in that some viruses use cellular enzymes during their life cycle and because viral infection can affect cellular regulation at the nucleic acid level. An understanding of the function of such regions and their extents of cross homology may help to classify types of virus–host interactions.

Simian virus 40 (SV40) is a convenient virus for such analysis because the entire nucleotide sequence is known (4–6), the genetic organization of the viral genome has been exhaustively studied, and, in many cases, the nucleotide sequence of regions of the genome can be related to a particular function [e.g., origin of replication, transcriptional starts and stops, coding regions, and large tumor antigen (T antigen) binding sites (reviewed in ref. 7)]. In this paper we report the isolation of cloned African green monkey (*Cercopithecus aethiops*) DNA segments that include regions of homology to the noncoding region surrounding the origin of replication of SV40.

MATERIALS AND METHODS

Preparation and Screening of Monkey DNA Library (8) in λ Phage. DNA of high molecular weight was isolated from the liver of an African green monkey (9). Portions of the DNA were subjected to partial cleavage by restriction endonuclease *Eco*RI (Boehringer) at various ratios of enzyme to DNA (0.01, 0.05, 0.1, 0.2, 0.5, and 1 unit/μg) for 15 and 30 min and then pooled. DNA fragments ranging in size between 14 and 20 kilobase pairs (kb) were then isolated by sucrose density gradient centrifugation (8). About 1 μg of the DNA was joined to 3 μg of isolated arms of *Eco*RI-cleaved λ Charon 4A DNA (10) with T4 DNA ligase and then packaged *in vitro* into infectious λ particles by the procedure of Enquist and Sternberg (11). More than 10^6 plaques were amplified, and 2.5×10^5 were screened (12) for hybridization to the *Hin*dIII C fragment of wild-type strain 776 SV40 (residues 5171–1046, map units 0.65 to 0.86 clockwise). [Throughout this paper we use the SV40 numbering system proposed by Buchman, Burnett, and Berg as described by Tooze (7) and corrected for the additional 17 bp recently reported (6); the total number of base pairs is 5243.] The *Hin*dIII C fragment was within recombinant plasmid pSVC1, which contains the *Hin*dIII C fragment inserted into the *Hin*dIII site of pBR322. It was kindly supplied by Cary Queen and was labeled with ^{32}P by nick translation (13). Hybridization and washing conditions for the nitrocellulose filters were as described below. All recombinant DNA experiments were carried out under the containment conditions specified by the National Institutes of Health *Guidelines for Recombinant DNA Research.*

Construction of Plasmids. DNA was isolated from λCaOri9 and cleaved with *Hin*dIII, and the resulting DNA fragments were ligated into the *Hin*dIII cleavage site of the plasmid pBR322 (14). This mixture of plasmid DNAs was used to transfect *Escherichia coli* strain RR1 (PRC no. 400 from the Plasmid Reference Center, Stanford University). Thirty-two tetracycline-sensitive (25 μg/ml) and ampicillin-resistant (25 μg/ml) transformants were arbitrarily selected, and plasmid DNA was isolated from each. Comparison of plasmid insert sizes with the fragment sizes resulting from *Hin*dIII digestion of λCaOri9 revealed that all internal *Hin*dIII fragments were represented approximately equally in this population.

Nucleotide Sequence Analysis. DNA fragments were dephosphorylated with alkaline phosphatase and 5′-end-labeled

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: SV40, simian virus 40; kb, 1000 base pair(s); bp, base pair(s); T antigen, large tumor antigen.

with [γ-^{32}P]ATP (Amersham, 2000 Ci/mmol; 1 Ci = 3.7×10^{10} becquerels) and phage T4 polynucleotide kinase (Boehringer). Sequence analysis was done by the procedure of Maxam and Gilbert (15).

Southern Blot Analysis. DNA fragments were separated by agarose gel electrophoresis after cleavage by the appropriate restriction endonuclease, transferred from the gel to nitrocellulose (16), and hybridized to a ^{32}P-labeled probe (13). The hybridization was performed in 0.45 M NaCl/0.045 M sodium citrate/0.2% bovine albumin/0.2% Ficoll/0.2% polyvinylpyrolidine/0.1% sodium lauryl sulfate containing 50 μg of denatured and sheared salmon sperm DNA, and 100 ng of denatured ^{32}P-labeled DNA probe (1×10^7 cpm) in a total volume of 30 ml for 16 hr at 67°C. Filters were then washed for three 30-min periods at 52°C in 0.03 M NaCl/3 mM sodium citrate/0.1% sodium laurel sulfate and then for three more 30-min periods at 52°C in 0.03 M NaCl/3 mM sodium citrate.

Spot tests were done in the same manner except that DNA (10–50 ng) was applied directly to the nitrocellulose filter. The filter containing DNA was sequentially wetted in 0.5 M NaCl/0.2 M NaOH, 0.5 M NaCl/0.2 M Tris, pH 7.4, and 1 M Tris, pH 7.4, and then baked at 80°C *in vacuo* for 2 hr. Hybridization was detected by autoradiography at −70°C with the use of intensifying screens.

RESULTS

Isolation, from Monkey Library, of λ Phage That Hybridize to SV40 *Hin*dIII C Fragment. The cloned segments of monkey DNA in the library contain varying numbers of internal *Eco*RI cleavage sites [0–20 sites (unpublished data)] and therefore reflect a fairly random distribution of *Eco*RI-generated fragments. Although the library represents a large proportion of the monkey DNA genome, the method by which it was prepared eliminates from representation those DNA segments that originate from genomic regions that contain widely spaced *Eco*RI cleavage sites (>20 kb).

λ phage were selected by hybridization to a ^{32}P-labeled plasmid which includes the SV40 origin of replication, T antigen binding sites, and the 5′ ends of both early and late mRNAs. With the assumptions that the haploid monkey genome is about $4–5 \times 10^9$ bp (17) and the average insert size is 1.7×10^4 bp, much of the library should be represented in 2.5×10^5 plaques. Six of 2.5×10^5 plaques were positive in both the initial and subsequent screenings. DNA was isolated from each and analyzed with restriction endonucleases. The six contained only three distinguishable inserts. Three phage, typified by λCaOri8 (Fig. 1*C*) contained one insert; two, typified by λCaOri9 (Fig. 1*E*) contained another; and one, λCaOri5 (Fig. 1*A*) contained a third (unpublished data).

Digests of DNA from each type of phage were analyzed by the Southern blot technique using ^{32}P-labeled SV40 DNA as a probe (Fig. 1 *B*, *D*, and *F*). The hybridization data show clearly that the SV40-like sequences are confined to limited regions of each monkey insert and confirm that each phage contains a distinct insert.

Further Analysis of λCaOri9. Digestion of λCaOri9 indicated the presence of multiple *Hin*dIII sites within the insert (three fragments <600 bp long were too light to be visible in Fig. 1*E*) and the fragment that hybridized with SV40 DNA was approximately 400 bp long (Fig. 1*F*). Further structural analysis of λCaOri9 was done by cloning the *Hin*dIII segments of the phage in pBR322 (Fig. 2*A*). Of the four different cloned fragments, only pCaOri9.32 hybridized to SV40 and it was, as expected, about 400 bp long. The order of cloned fragments within λCaOri9 was determined by Southern blot analysis of cleaved λCaOri9 with radioactively labeled plasmids as hybridization probes (summarized in Fig. 2*A*).

Homology to SV40 in λCaOri9 is Limited to Region Around Origin of Replication. Results of hybridizations of ^{32}P-labeled λCaOri9 or pCaOri9.32 to blots containing *Hin*dIII-restricted SV40 DNA were identical and were consistent with the fact that the phage contains no SV40-like sequences other than those in the *Hin*dIII C fragment (Fig. 3*A*). Further delimitation of the

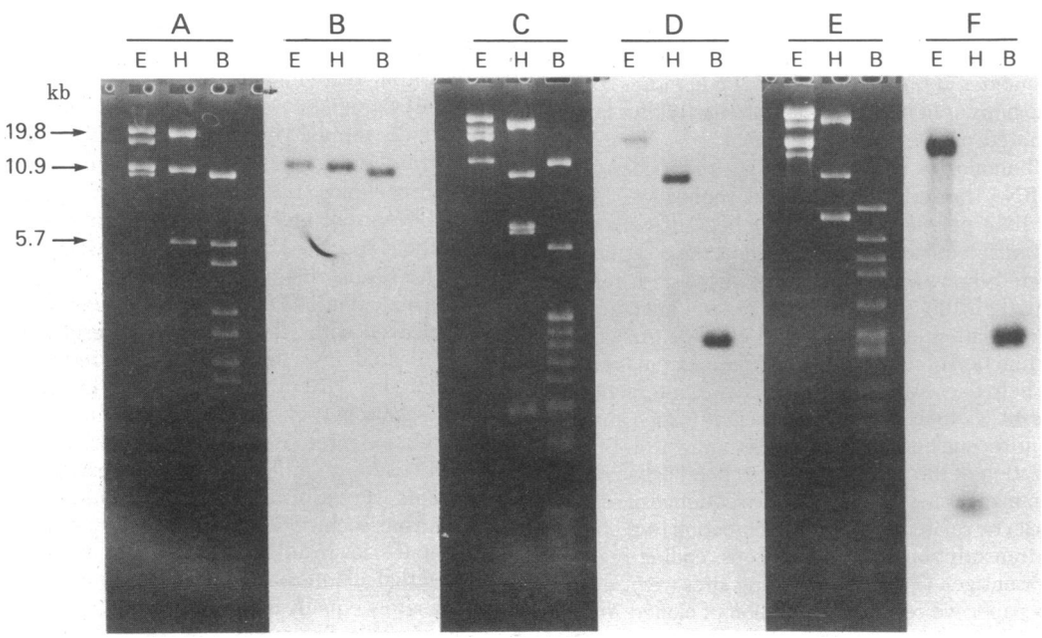

FIG. 1. A 0.5-μg sample of each purified λ phage was digested with the indicated restriction endonuclease and electrophoresed at 2 V/cm for 20 hr through a 0.8% agarose gel. DNA was visualized by fluorescence enhancement with ethidium bromide [*A*(λCaOri5), *C*(λCaOri8), and *E*(λCaOri9)], transferred to nitrocellulose paper, and hybridized with ^{32}P-labeled SV40 DNA (*B*, *D*, and *F*). Arrows point to fragments derived from the λ vector (the slowest moving, unmarked band in the *Eco*RI channels is hybridized λ arms). Abbreviations: E, *Eco*RI; H, *Hin*dIII; B, *Bgl* I.

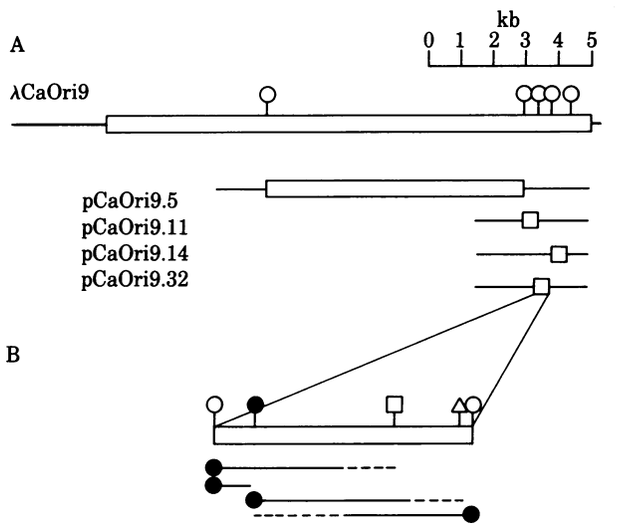

FIG. 2. (*A*) Diagram of λCaOri9, showing sites cleaved by *Hin*dIII and *Hin*dIII-generated fragments subcloned into pBR322 (pCaOri9.5, -9.11, -9.32, and -9.14). Bar, monkey sequences; line, pBR322 sequences. (*B*) Diagram of pCaOri9.32, showing the strategy used for primary nucleotide sequence determination. Restriction endonuclease sites used: ♀, *Hin*dIII; ♦, *Alu* I; ⍒, *Hpa* II; ⇑, *Ava* II. Each line below shows a fragment from which primary nucleotide sequence data was obtained. The full portion of each line is the portion in which the sequencing results were clearly interpretable. The solid circle on each line shows the end of the fragment labeled with [γ-^{32}P]ATP.

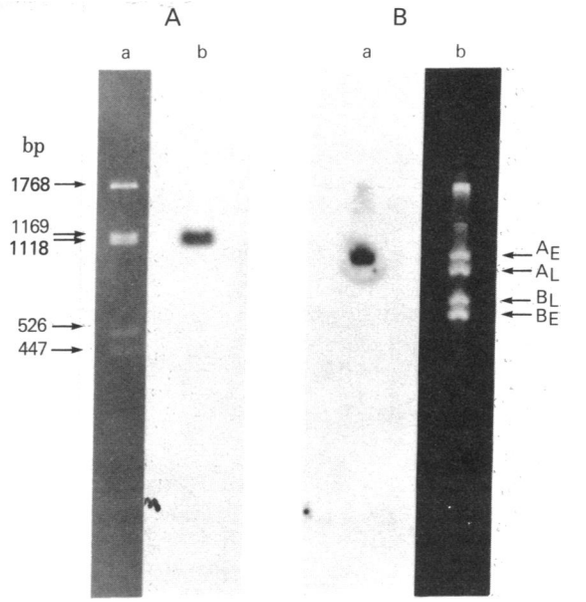

FIG. 3. (*A*) A 0.5-μg sample of SV40 DNA was digested with *Hin*dIII and electrophoresed at 2 V/cm through a 2% agarose (wt/vol) gel for 15 hr. DNA was visualized by fluorescence enhancement with ethidium bromide (lane a), transferred to nitrocellulose paper, and hybridized with ^{32}P-labeled pCaOri9.32 (lane b). (*B*) A 1-μg sample of SV40 DNA was digested with both *Hpa* II and *Bam*HI. The DNA was denatured in 0.2 M NaOH and electrophoresed at 2 V/cm through 2% agarose for 15 hr. DNA was visualized as above (lane b), transferred to nitrocellulose paper, and hybridized with ^{32}P-labeled pCaOri9.32 (lane a). The fragments were identified as the early- or late-coding strands of SV40 as described (18).

bounds of SV40 homology was accomplished by Southern analysis of SV40 DNA that had been cleaved with both *Hpa* II (residue 346) and *Bam*HI (residue 2533), treated with NaOH to separate strands, and electrophoresed (Fig. 3*B*). Hybridization of ^{32}P-labeled pCaOri9.32 to the A fragment indicates that homology to the *Hin*dIII C portion of SV40 does not extend beyond the *Hpa* II site at 346. As we expected from preliminary experiments, the probe hybridized only with the A_E fragment and not with the A_L. It may be that one strand of the ^{32}P-labeled pCaOri9.32 probe was unavailable for hybridization. This situation could be explained by extensive secondary structure resulting from the very high guanosine content (see below) of the strand that would be expected to hybridize with A_L. It is in-

```
        10          20          30          40          50          60
AGCTTGACTA GGGGCACGGC GGGCGCCTCG CCTGGCGAGG GGCGGGGCGG ACGGGGCGGC

        70          80          90          100         110         120
CAGCTAGGGG CGGAGCAGGG CGGAGCAGGG CGGGTGCAAG GCGAGTGGTG GGGACCCAGT

        130         140         150         160         170         180
TGGAAGCGGA GGCCATAGGG TAAGCCCGCC GAGGGGCGGC GGGGCGGCAG GGGAGGCCTG

        190         200         210         220         230         240
ATGACGACCT GCGCGTGCGG CTCGGCCTGA AGGGGCAAGC AGAGGACTTG GCAGGGCTGG

        250         260         270         280         290         300
TGAGGGGCGG GGCAGGCGGC GCGGTGGGGG CGGGCCGAGC CCGGAGGTCG GATGAGCGGA

        310         320         330         340         350         360
CACGGCCCGA CGCGCGAAGC CATGCAAGTA GGTGGCTCCC GACGGCCCCG CTTGAATTTC

        370         380
GATCCCAAAC CGGGTCCXXX CTCCCGGCC
```

FIG. 4. Nucleotide sequence of the monkey DNA insert in pCaOri9.32. The strategy used for sequence determination is shown in Fig. 2*B*. The underlined sequences are homologous to segments within the region of dyad symmetry in SV40 DNA. The overlined sequences are homologous to the short repeated segments that occur to the late side of the dyad symmetry in SV40 DNA. The broken underlines indicate sequences that are homologous to a portion of the 72-bp repeated sequence in SV40 DNA. The arrowheads to the right indicate homology in the pCaOri9.32 strand as written; arrowheads to the left indicate homology in the other strand [always compared to the nucleotide sequence of the SV40 DNA strand corresponding to the sense of late RNA (7)]. Homology was investigated by using a computer program (19). Although many homologies were detected, only those at least 10 bp long and more than 82% matched are marked. We note that the *Ava* II site predicted for residues 112–116 could not be confirmed by direct digestion; it is likely that the A at 114 is incorrect.

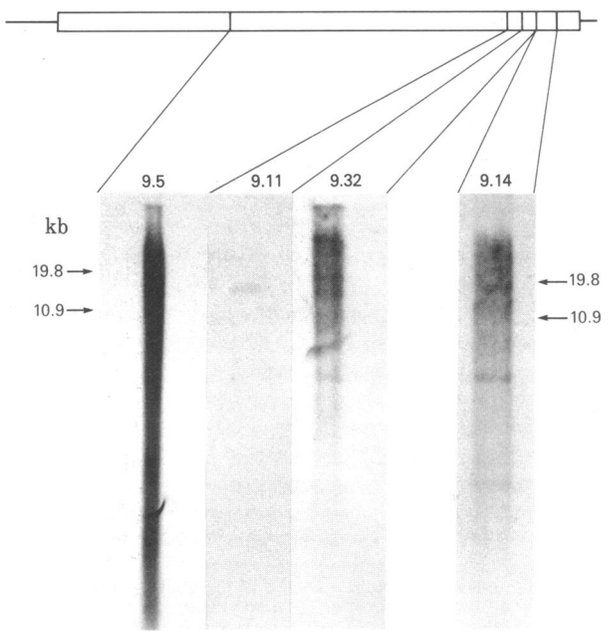

FIG. 5. DNA fragments from total monkey DNA that are homologous to the monkey DNA cloned in λCaOri9. Twenty micrograms of monkey DNA was digested with *Eco*RI and electrophoresed at 2 V/cm through a 0.8% agarose gel for 20 hr. The DNA was transferred to nitrocellulose paper and hybridized to the indicated plasmid labeled with ^{32}P by nick translation. All the probes contained about 10^8 cpm/μg but exposure times for the autoradiograms varied: pCaOri9.5, 2 hr; pCaOri9.11, -9.32, and -9.14, 10 days. Arrows point to fragments derived from the λ vector.

teresting to consider that this structural feature may be functionally significant.

Primary nucleotide sequence analysis of the monkey insert in pCaOri9.32 (Fig. 4) shows that the significant homology is dispersed over about 325 bp of the insert and is limited to a region between residues 5213 and 250 of SV40. The homology is complex and is analyzed further in the *Discussion*. However, we point out here several interesting features of the sequence itself. G+C constitutes 74.7% of the bases [total monkey DNA is about 40% G+C (20)]. The Gs and Cs are distributed asymmetrically: the strand shown in Fig. 4 is 47.2% G. Also, 11.5% of the dinucleotides are the usually rare C-Gs. There are many repeated sequences within the segment. Among the more striking are 10 repeats of G-G-G-C-G-G, and 2 of G-A-G-G-G-G-C-G-G-G-G-C (residues 37–48 and 242–253). There is also a 16-bp overlapping repeat from residues 68 to 83 and 78 to 93, a configuration we term a "canon."

Organization of λCaOri9 Monkey Sequences in the Monkey Genome. The occurrence of genomic DNA homologous to the various subcloned segments of λCaOri9 was investigated by Southern blot analysis of total monkey DNA digested exhaustively with *Eco*RI (Fig. 5). Autoradiograms of blots hybridized with pCaOri9.5 revealed extensive hybridization throughout all size ranges indicating the presence of a highly repeated sequence interspersed among many different size classes of *Eco*RI fragments. Separate experiments demonstrated that this sequence is not the very highly repeated α component of monkey DNA (20, 21). pCaOri9.11 hybridized only with a fragment about 16 kb long that comigrated with the monkey insert in λCaOri9. From the extent of this hybridization we conclude that the total monkey segment in λCaOri9 probably occurs only once or twice. In contrast to experiments with pCaOri9.11, Southern blots using pCaOri9.32 and pCaOri9.14 as probes revealed a background of hybridization as well as several clearly defined bands (not all easily visible on the photograph). From the extent of hybridization, we estimate that each band represents multiple genomic copies and that the background may represent additional related sequences.

Spot tests were carried out in which λCaOri8 and λCaOri5 were hybridized with ^{32}P-labeled pCaOri9.5, pCaOri9.14, pCaOri9.11, and pCaOri9.32 (not shown). pCaOri9.5 and pCaOri9.14 hybridized with both phage; pCaOri9.32, which contains the entire SV40 homology in λCaOri9, hybridized lightly to the other phage, suggesting only limited homology. pCaOri9.11 did not hybridize with either of the other two phage. These data are consistent with the conclusions reached from the experiments with genomic blots and also indicate that the SV40-like segments in all three phage have some common near neighbors.

Southern analysis of *Eco*RI-cleaved monkey DNA was also attempted with pSVC1 as a probe but autoradiography of these blots revealed only light hybridization and no distinct banding pattern. It was not possible, therefore, to determine whether all size fragments that hybridize to pCaOri9.32 also contain SV40 sequence homology.

DISCUSSION

We have isolated three different cloned segments of the monkey genome that hybridize under stringent conditions to the

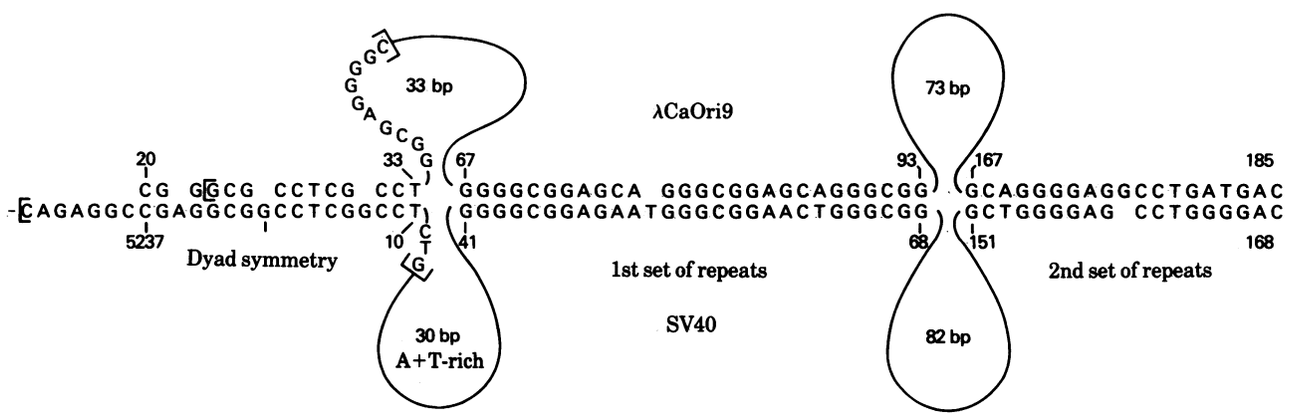

FIG. 6. Comparison of sequences around the SV40 origin of replication with those in pCaOri9.32. Regions of dyad symmetry are bracketed. Numbering systems are according to ref. 7 for SV40 and as in Fig. 4 for pCaOri9.32. This is one of a number of possible ways to match the sequences. Regions unmatched in this display are looped out but also contain some regions of homology.

area of the SV40 genome surrounding the origin of replication. Analysis of plaque hybridization data indicates that these three segments of monkey DNA represent the closest homology to the SV40 origin region to be found in the monkey library.

We have characterized one of these segments in some detail. Hybridization data and sequence analysis of the monkey segment cloned in λCaOri9 show that homology is restricted to between residues 5213 and 250 of SV40 DNA. Within SV40, this 280-bp-long region is involved in control rather than in coding functions. It contains sequences corresponding to the origin of DNA replication (reviewed in ref. 7), the putative promotors for early mRNA synthesis (22, 23), the 5′ ends of early mRNAs (24, 25), the 5′ ends of some 19S late mRNAs (see ref. 7), and binding sites for T antigen (7, 26, 27). Each of these control functions is thought to be directly related to several distinctive regions of DNA sequence. In a clockwise direction, the pertinent SV40 sequences include: two sets of dyad symmetries of 15 and 27 residues, respectively, an A+T-rich region of 20 residues, and two sets of repeated sequences (Fig. 6). Furthermore, in SV40 minichromosomes this region is preferentially sensitive to nuclease, presumably reflecting an unusual chromatin configuration imposed by the nucleotide sequence (28, 29). As shown in Fig. 4, the homologous segment of monkey DNA has five different regions of homology (>82%) to the SV40 area of dyad symmetry (SV40 residues 5213–15), no homology to the SV40 A+T-rich region, six regions of homology (>82%) to the first set of repeated sequences (SV40 residues 39–103), and two regions of homology to the second set of repeats (SV40 residues 107–250).

Although the total homology to three separate small regions of SV40 is considerable, the arrangement and number of copies of each are quite different. In an attempt to mimic the structural arrangement in SV40 we have chosen to display the monkey DNA segment as shown in Fig. 6 although, by looping out various other regions of SV40 and monkey DNA, various different homologies can be selected. The particular similarities emphasized in Fig. 6 were chosen for several reasons; the region of homology to the 27-bp dyad symmetry of SV40 includes a 22-bp region of dyad symmetry in the monkey DNA, the distance between the dyad symmetry and the first set of repeats is approximately equal to that found in SV40, the distance between the first set of repeats and the second set of repeats is again approximately equal to that found in SV40, and the sequence G-G-G-X-G-G-A-G—which is common to (and often repeated within) the origin regions of SV40, BK virus, polyoma, simian adenovirus 7, and various human adenoviruses (reviewed in refs. 30 and 31)—is repeated.

Nevertheless, the region of dyad symmetry and the two sets of repeats all may play separate as well as overlapping roles in control of SV40 functions and therefore the significance of their relative juxtaposition in SV40 may be specific to the virus. If, in the monkey genome, the homologous regions are also involved in control functions, then their scrambled order may be dictated by different functional requirements. This view suggests that control units may be constructed in modular fashion. Although most of the SV40 genome may have evolved from an earlier papovavirus prototype (32), the clustered regulatory signals to the late side of the origin more likely reflect adaptations for efficient production of the virus in monkey cells (33). Such adaptation may have involved the coopting of modular units of the monkey genome. Further consideration of these points awaits structural analysis of the remaining SV40-like monkey cloned segments; the weak hybridization of pCaOri9.32 to λCaOri8 and λCaOri5 suggests that these other monkey DNA segments have alternate arrangements of the SV40 origin-like sequences. Experiments testing the possible functions of the sequence in replication, transcription, and T antigen binding are also required.

We thank our colleagues Cary Queen, Ronald E. Thayer, and Francine Eden for generous help and advice. May Liu prepared the manuscript with patience and skill.

1. Stekelin, D., Varmus, H. E., Bishop, J. M. & Vogt, P. K. (1976) *Nature (London)* **260**, 170–173.
2. Spector, D. H., Smith, K., Padgett, T., McCombe, P., Roulland-Dussoix, D., Mascovici, C., Varmus, H. E. & Bishop, J. M. (1978) *Cell* **13**, 371–379.
3. Frolova, E. I. & Georgiev, G. P. (1979) *Nucleic Acids Res.* **7**, 1419–1428.
4. Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Van de Voorde, A., Van Heuverswyn, H., Van Herreweghe, J., Volckaert, G. & Ysebaert, M. (1978) *Nature (London)* **273**, 113–120.
5. Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Ghosh, P. K., Celma, M. L. & Weissman, S. M. (1978) *Science* **200**, 494–502.
6. Van Heuverswyn, H. & Fiers, W. (1979) *Eur. J. Biochem.* **100**, 51–60.
7. Tooze, J. (1980) *DNA Tumor Viruses*, Molecular Biology of Tumor Viruses, Part 2 (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY), 2nd Ed.
8. Maniatis, T., Hardison, R. C., Lacy, E., Lauer, J., O'Connell, C., Quon, D., Sim, G. K. & Efstratiadis, A. (1978) *Cell* **15**, 687–701.
9. Britten, R. J., Graham, D. E. & Neufeld, B. R. (1974) *Methods Enzymol.* **29E**, 363–418.
10. Williams, B. G. & Blattner, F. R. (1979) *J. Virol.* **29**, 555–575.
11. Enquist, L. & Sternberg, N. (1980) *Methods Enzymol.* **68**, 281–298.
12. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
13. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. (1977) *J. Mol. Biol.* **113**, 237–251.
14. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene* **2**, 95–113.
15. Maxam, A. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
16. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
17. Singer, D. S. (1979) *J. Biol. Chem.* **254**, 5506–5514.
18. Ferdinand, F.-J., Brown, M. & Khoury, G. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5443–5447.
19. Queen, C. & Korn, L. J. (1980) *Methods Enzymol.* **65**, 595–609.
20. Maio, J. J. (1971) *J. Mol. Biol.* **56**, 579–595.
21. Rosenberg, H., Singer, M. & Rosenberg, M. (1978) *Science* **200**, 394–402.
22. Benoist, C. & Chambon, P. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3865–3869.
23. Gluzman, Y., Sambrook, J. F. & Frisque, R. J. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3898–3902.
24. Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M. & Weissman, S. M. (1979) *J. Virol.* **30**, 279–296.
25. Thompson, J. A., Radonovich, M. F. & Salzman, N. P. (1979) *J. Virol.* **31**, 437–446.
26. Tjian, R. (1978) *Cell* **13**, 165–179.
27. Shalloway, D., Kleinberger, T. & Livingston, D. M. (1980) *Cell* **20**, 411–422.
28. Varshavsky, A. J., Sundin, O. & Bohn, M. (1979) *Cell* **16**, 453–466.
29. Wigmore, D. J., Eaton, R. W. & Scott, W. A. (1980) *Virology* **104**, 462–473.
30. Seif, I., Khoury, G. & Dhar, R. (1979) *Cell* **18**, 963–977.
31. Tolun, A., Alestrom, P. & Pettersson, U. (1979) *Cell* **17**, 705–713.
32. Soeda, E., Maruyama, T., Arrand, J. R. & Griffin, B. E. (1980) *Nature (London)* **285**, 165–167.
33. Yang, R. C. A. & Wu, R. (1979) *Nucleic Acids Res.* **7**, 651–668.