MOLECULAR AND CELLULAR BIOLOGY, Nov. 1983, p. 1930–1936
0270-7306/83/111930-07$02.00/0

Vol. 3, No. 11

Naturally Arising Recombinants That Are Missing Portions of the Simian Virus 40 Regulatory Region

MARY WOODWORTH-GUTAI,* ANTHONY CELESTE, LOWELL SHEFLIN, AND MORTON SCLAIR
Department of Cell and Tumor Biology, Roswell Park Memorial Institute, Buffalo, New York 14263

Received 25 April 1983/Accepted 12 August 1983

When simian virus 40 (SV40) is serially passaged at high multiplicity, a heterogeneous collection of naturally arising variants is generated. Those which are the most abundant presumably have a selective replicative advantage over other defective and wild-type helper SV40s. Two such naturally arising host-substituted variants of SV40 have been characterized in terms of complete nucleotide sequence determination. Evolutionary variant *ev*-1101 (previously isolated by Lee et al., Virology **66**:53–69, 1975) is from undiluted serial passage 13, whereas *ev*-2101 is newly isolated from undiluted serial passage 6 of an independently-derived evolutionary series. Both variants contain a five-times tandemly repeated segment of DNA consisting of viral *Hin* C and *Hin* A sequences that have recombined with a segment of host DNA that is not highly reiterated in the monkey genome. The monkey segment differs in the two variants as does the size of the viral segment retained. In two additional host-substituted variants, *ev*-1102 (previously isolated from serial passage 20 by Brockman et al., Virology **54**:384–397, 1973) and *ev*-1108 (newly isolated from serial passage 40), the SV40 sequences derived from the replication origin are present as inverted repetitions. The inverted repeat regions of these two variants have been analyzed at the nucleotide sequence level and are compared with SV40 variant *ev*-1104 from passage 45 (previously characterized by Gutai and Nathans, J. Mol. Biol. **126**:259–274, 1978). The viral segment containing the regulatory signals for replication and viral gene expression is considerably shortened in later serial passages as demonstrated by these five variants. It is of interest that the variants presumably arose due to their enhanced replication efficiency, yet are missing some of the sequence elements implicated in the regulation of replication. Furthermore, a comparison of the structure of the replication origin regions indicates that additional changes occur in the SV40 regulatory region with continued undiluted serial passage.

After undiluted serial passage of simian virus 40 (SV40) at a high multiplicity of infection, variants evolve in which the viral DNA is rearranged due to recombination events that have deleted and duplicated parts of the SV40 genome (3). In addition, variants evolve in which host cell DNA has recombined with the viral genome. Presumably these variants evolve because they have a selective replicative advantage over other defective and wild-type SV40s. All evolutionary variants isolated to date have at least two copies of the *ori*-containing region of the SV40 genome, and generally four to six copies of the origin of DNA replication are present in later-passage variants (11, 15, 23, 30; Woodworth-Gutai, unpublished data). These variants have evolved by unknown recombination events in accordance with undefined cellular and viral constraints. However, since the viral *ori* signal is the only known *cis*-acting element required for propogation, other viral sequences are presumably dispensible and can be lost through recombination and selection.

One approach to studying recombination events in SV40-infected cells is to study the removal of nonfunctional sequences during the evolution of naturally arising variants. By this approach we should also gain insight into the functional *cis*-acting signal for viral DNA replication. This study describes the characterization of a newly isolated variant, *ev*-2101, from undiluted serial passage 6 and a previously isolated variant, *ev*-1101 (15), from passage 13 of an independently derived evolutionary series. Also included is a comparison of the SV40 replication origin sequences present as inverted repeat regions in three variants: *ev*-1102 (4), previously isolated from passage 20, *ev*-1108, newly isolated from passage 40, and *ev*-1104 (11), previously characterized from passage 45. The structure of

these variants is compared with those of variants from earlier and later passages in terms of recombination events and retention of functional viral sequence.

MATERIALS AND METHODS

SV40 (small plaque, strain 776) was grown on BSC-1 cells, a line of African green monkey kidney cells. Variant DNA was isolated as detailed previously (33). Serial passage of SV40 in BSC-1 cells and cloning of host-substituted variants have been described previously (4, 5). Evolutionary variants *ev*-1101, *ev*-1102, and *ev*-1104 were previously cloned from undiluted serial passages 13, 20, and 45, respectively (4, 11, 15). Restriction enzymes (New England BioLabs) were used as recommended by the supplier. Dephosphorylation with bacterial alkaline phosphatase (Worthington Diagnostics) and 5'-end labeling with [γ-^{32}P]ATP (1,000 to 3,000 Ci/mmol; New England Nuclear Corp) and polynucleotide kinase (P-L Biochemicals) have been described previously (33). DNA restriction fragments from which one 5'-labeled terminus was removed by secondary restriction cleavage or strand separation were sequenced by the dimethyl sulfate-hydrazinolysis method of Maxam and Gilbert (16). Products were fractionated on 0.4-mm thin gels (21) composed of 8, 10, or 20% acrylamide–8 M urea.

RESULTS

Characterization of variants. Uniformly labeled and end-labeled restriction fragments as well as double-digest products were utilized to construct restriction maps. When the *ev*-2101 genome was digested with *Bgl*I or *Hin*dIII, a monomer fragment 16% the size of wild-type SV40 DNA was produced. Since the size of the variant genome is 79% of full-length SV40 DNA by electrophoretic mobility the 16% monomer unit must be tandemly repeated in the variant genome a total of five times. By electron microscopy, Lee et al. (15) estimated that the size of *ev*-1101 was 84.8 ± 4.7% and concluded that the monomer 17% fragment is tandemly repeated five times in the variant genome. The inverted repeat regions of *ev*-1102 and *ev*-1108 were visualized by electron microscopy of single-stranded molecules generated from form II molecules which had been denatured and renatured. Single-stranded circles of variant *ev*-1102 have three symmetrically distributed duplex segments, i.e., three inverted repeat sequences, whereas *ev*-1108 has three asymmetrically placed inverted repeat segments. Restriction maps of the inverted repeat regions were constructed before nucleotide sequence analysis.

Nucleotide sequencing analysis. The nucleotide sequence analysis of overlapping restriction fragments and separated complementary strands confirmed the restriction endonuclease mapping studies. The diagrammatic composite of the nucleotide sequencing data is summarized in Fig. 1a and b. The viral DNA sequences in each tandemly repeated monomer unit of *ev*-2101 total 295 base pairs (bp). The viral DNA sequences begin at nucleotide position 0126 in wild-type SV40 *Hin* C and extend to position 5075 in wild-type SV40 *Hin* A (Fig. 1a). Recombination with host DNA sequences occurred at these two viral DNA sites. The nucleotide sequence at the two viral-host junctions is shown in Fig. 2a and b. The viral *Hin* A and *Hin* C segments spanning the two recombinant joints have an adenine + thymine (A+T) content that is average to poor (59 and 43%, respectively). In contrast, the 50-bp portion of adjacent host DNA has an A+T base composition of 70% at the *Hin* A-host junction and 68% A+T at the *Hin* C-host junction. There is a striking cluster of 17 adenines in the host sequence and a 14-bp A+T cluster in the viral *Hin* A sequence. Adjacent to the viral-host junctions are short sequences of 4 or 5 bp of shared homology. The total 529-bp host DNA sequence present in the monomeric unit of *ev*-2101 is shown in Fig. 3.

The 864-bp monomer repeat unit of *ev*-1101 contains a 300-bp segment of host DNA which had recombined with a 564-bp viral segment at nucleotide positions 0066 of *Hin* C and 4746 of *Hin* A (Fig. 1a). The nucleotide sequence of these two viral-host junctions is shown in Fig. 2c and d. The base composition of the sequences adjacent to the viral *Hin* A-host junction is 72% A+T (host) and 62% A+T (viral). On the other hand, the host segment at the *Hin* C-host junction is strikingly A+T poor (40%), as is its viral counterpart (44%). There is a 15-bp stretch of A+T in the viral *Hin* C segment, but like the small 4 to 6-bp stretches of shared homology, it is not near the crossover. The complete host DNA sequence in shown in Fig. 4.

The *ori*-containing inverted repeat structures of *ev*-1102 and *ev*-1108 are compared with that of the previously reported *ev*-1104 (11) in Fig. 1b. The passage 20 inverted repeat region contains no host sequences and simply consists of two copies of the *ori*-containing viral *Hin* C sequence: one inverted relative to the other, and one containing 149 bp more *Hin* C sequence than the other. At passage 40, in addition to two copies of the *ori*-containing viral *Hin* C DNA with one copy inverted relative to the other, there is a 69-bp host DNA insert between the viral segments and an insertion of a 4-bp duplication of *Hin* C sequence in one arm of the inverted repeat. Appearing even more complex is the passage 45 inverted repeat segment which has retained the same 69-bp host insert between the inverted *ori*-containing viral segments and in addition has a 6-bp duplication of *Hin* C sequence in one arm of the inverted repeat and a larger 30-bp duplication (consisting of 8 bp of

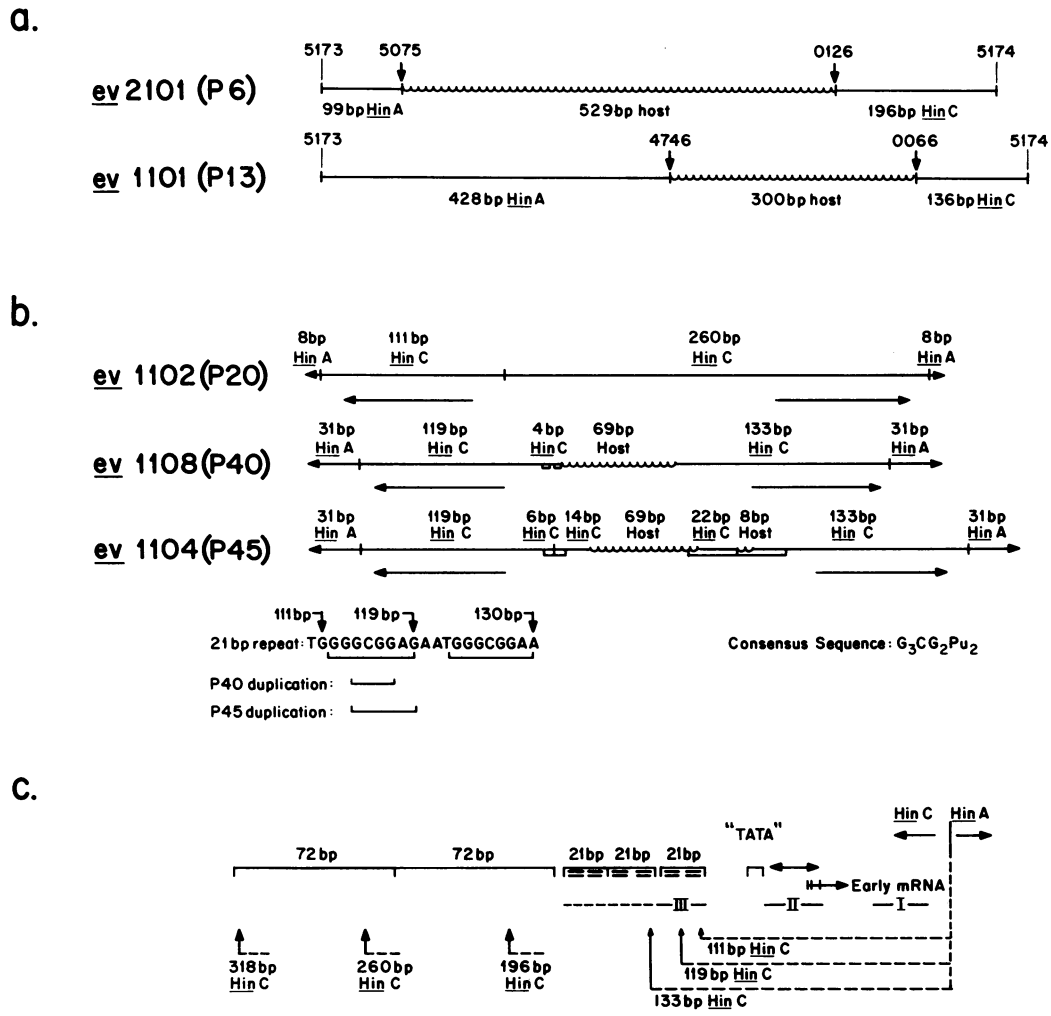

FIG. 1. Diagrammatic representation of the viral and cellular organization of five evolutionary variants and a comparison of their *Hin* C sequence content. Numbering of SV40 nucleotides corresponds to the SV numbering system in reference 29. P refers to serial passage number. The viral sequence is indicated by a solid line, and the host sequence is indicated by a wavy line. (a) The monomeric unit (repeated five times) of *ev*-2101 (P6) and *ev*-1101 (P13). (b) The inverted repeat region (three copies per variant) of *v*-1102 (P20), *ev*-1108 (P40), and *ev*-1104 (P45). Horizontal arrows denote the inverted orientation of the *Hin* C segments. A sequence duplication is indicated by two brackets. (c) The regulatory region of SV40 and the locations of the 111-, 119-, 133-, 196-, and 260-bp *Hin* C segments relative to the 21-bp repeats. Roman numerals indicate the three T antigen binding sites (28). Location of two copies of the consensus sequence in each 21-bp repeat is depicted by a double dash (=). Also indicated are the TATA box and the direction, origin, and cap sites of early mRNA. The 27-bp sequence with a perfect twofold axis of symmetry is indicated by a double-headed arrow (↔).

host DNA and 22 bp of viral *Hin* C DNA) in the other arm.

DISCUSSION

In this study, we have completely characterized the variant genomes of *ev*-2101 (passage 6) and *ev*-1101 (passage 13) at the nucleotide sequence level and also analyzed variations which occur in the nucleotide sequence of *ori*-containing segments with continued serial passage, as exemplified by variants *ev*-1102 (passage 20), *ev*-1108 (passage 40), and *ev*-1104 (passage 45). The passage 6 variant has retained more of the viral *Hin* C *ori*-containing sequence than has the passage 13 variant. The progression toward decreasing amounts of viral *Hin* C sequence with prolonged serial passage is nicely demonstrated by comparing the extent of viral *Hin* C sequence retained in passage 3 variants (543 bp for *ev*-1119, 528 bp and 305 bp for *ev*-1114, and 543 bp

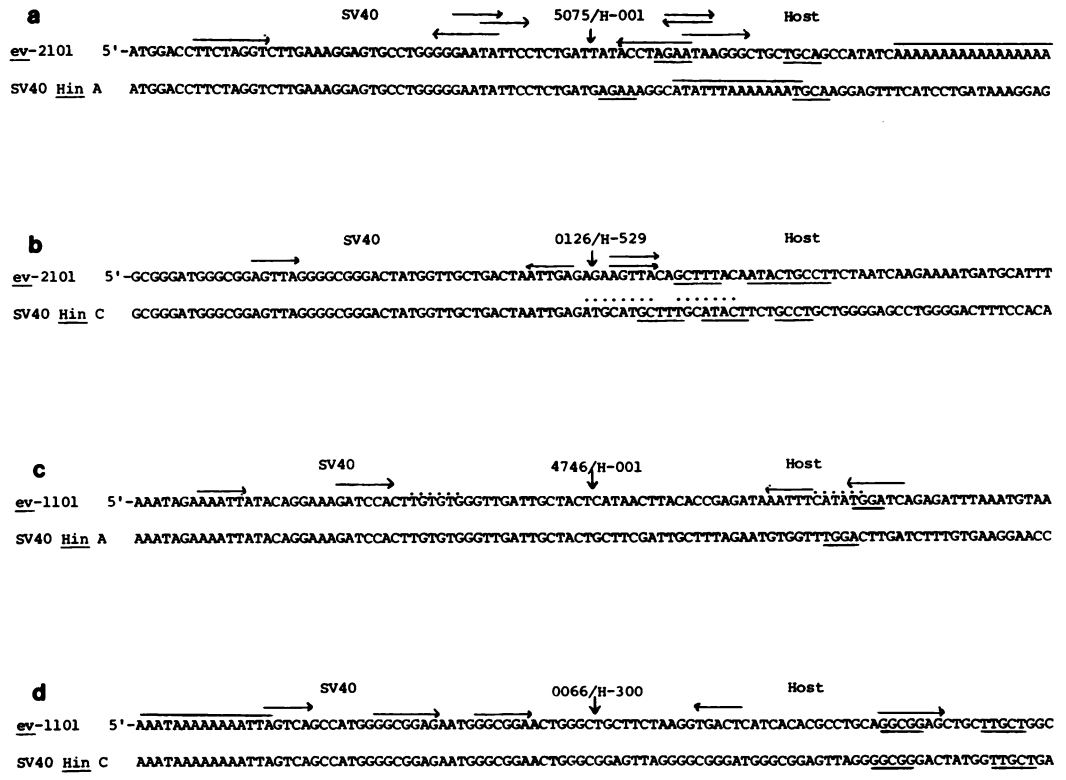

FIG. 2. Nucleotide sequence of the viral-host junctions in *ev*-2101 (a and b) and *ev*-1101 (c and d). The vertical arrow locates the junction between viral and host DNAs. On the line directly below the recombinant sequence, the viral sequence is continued beyond the point of crossover for comparative purposes. Stretches of homology shared by the parental recombining DNAs are underlined; interesting clusters of A+T bases are overlined with a solid line, and alternating purine-pyrimidine bases are overlined with dots. Direct repeats are indicated by head-to-tail arrows (→ →); inverted repeats are indicated by head-to-head arrows (→ ←); and true palindromes are indicated by tail-to-tail arrows (← →).

and 298 bp for *ev*-1117 [33]) with passage 5 (355 bp for *ev*-1110, 353 bp for *ev*-2102, and 346 bp for *ev*-2114 [unpublished data]), with passage 6 (196 bp for *ev*-2101), and with passage 13 variants (139 bp for *ev*-1103, 136 bp for *ev*-1101 [12, 17]). This trimming of viral *Hin* C sequence with additional undiluted serial passage is not seen in terms of the extent of viral *Hin* A sequences retained. For example, passage 5 variants *ev*-1110, *ev*-2102, and *ev*-2114 have 66, 212, and 4 bp of *Hin* A, respectively; *ev*-2101 (passage 6) has 99 bp of *Hin* A; and two variants from passage 13, *ev*-1101 and *ev*-1103, have 428 and 31 bp of *Hin* A, respectively. There appear to be different constraints on the recombination events that trim the viral *Hin* C sequence to approximately 136 bp, whereas the amount of *Hin* A sequence fluctuates as evolution of the variant progresses. Since the *cis*-acting sequences within *Hin* C are required for replication of the evolutionary variant, these naturally arising recombinants should provide some insight into the identity of the functional sequences that are essential for efficient replication. The bidirectional origin of SV40 DNA replication was first mapped at about 0.67 map unit in the *Hin* C fragment of the SV40 physical map (6, 9). The analysis of evolutionary variants and deletion mutants has narrowed the limits of the origin signal for SV40 DNA replication to a segment of approximately 65 bp from about 0.665 to 0.675 map unit (8, 11, 27). It has also been shown that single nucleotide changes or small deletions within this segment abolish or alter replication (10, 18, 22). According to Hay and DePamphilis (13), the in vivo initiation sites for DNA synthesis map both inside and outside the 65-bp *ori* sequence, but the single most frequently used site is located within the *ori* segment in the 27-bp sequence possessing a perfect twofold rotational axis of symmetry. Those initiation sites that map outside include some preferred sites within the three 21-bp, quanine-plus-cytosine (G+C)-rich tandem re-

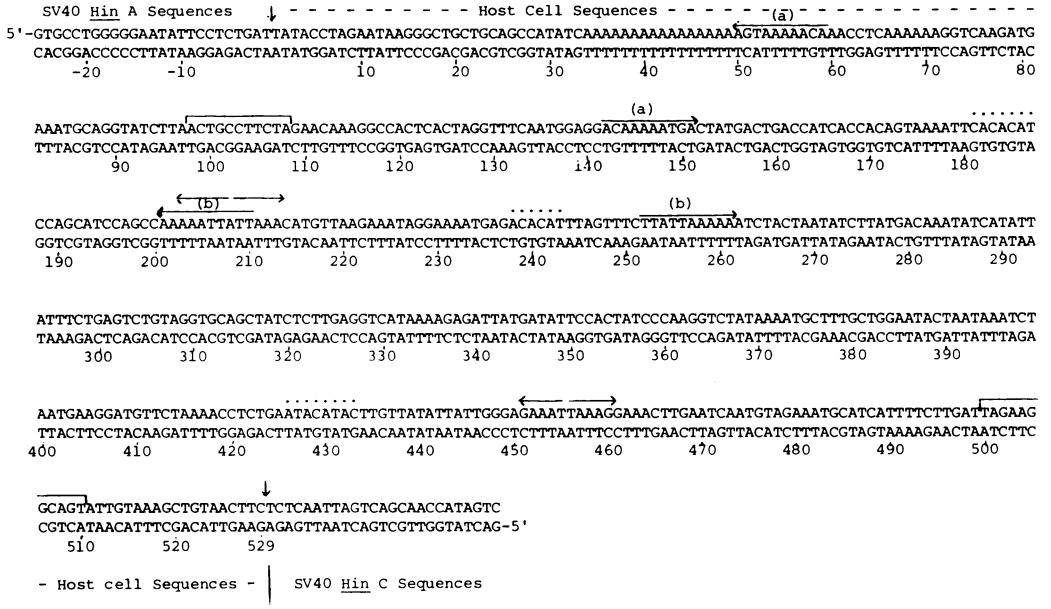

FIG. 3. Complete nucleotide sequence of the 529-bp segment of host DNA. Included are the 25-bp portions on either end of the viral segment covalently linked to host DNA in the repeating unit of *ev*-2101. True palindromes are designated by tail-to-tail arrows; corresponding ones that are separated by a large sequence interval are indicated by (a) or (b). The bracketed sequence at nucleotide residues 500 to 510 is the inverted repeat of the sequence located at residues 98 to 108. Stretches of alternating purines and pyrimidines are overlined with dots.

peats (Fig. 1c) and may recognize the consensus sequence $G_3CG_2Pu_2$, which is repeated six times within the 21-bp triplet. That the G+C-rich repeats outside the *ori* sequence may play a regulatory role in replication has also been suggested by Bergsma et al. (1), who report that the loss of two out of three of the 21-bp repeats drops replication efficiency at least twofold.

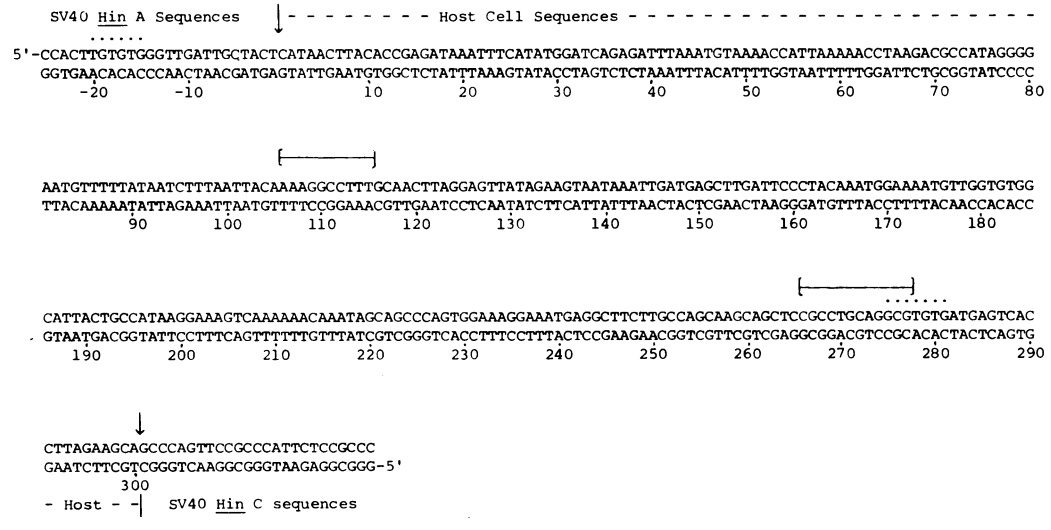

FIG. 4. Complete nucleotide sequence of the 300-bp segment of host DNA incorporated into the repeating unit of *ev*-1101. The interspersion of host DNA and viral DNA is demonstrated by inclusion of 25 bp of the viral sequences flanking the host DNA. A bracketed line locates sequences with a twofold axis of symmetry, whereas dots designate sequences with alternating purines and pyrimidines.

Although *ev*-2101 (passage 6) has retained all three 21-bp G+C-rich tandem repeats, later-passage variants *ev*-1101, *ev*-1102, *ev*-1104, and *ev*-1108 have retained only one intact copy of the 21-bp G+C-rich repeat, yet these naturally arising variants are the dominant species presumably because they have a replicative advantage. There are at least two possibilities: (i) multiple copies of *ori* may be better than more than one copy of the 21-bp G+C-rich repeat; or (ii) the host sequence in these variants may be able to effectively substitute for the 21-bp G+C-rich putative regulatory sequence. For example, in the case of *ev*-1101 a seven-nucleotide sequence, GGCGGAG (which is colinear with seven of eight nucleotides of the consensus sequence), is present in the host sequence at 32 bp from the viral-host junction (Fig. 2d). Furthermore, variants containing inverted repetitions of the SV40 origin of replication are the predominant species after prolonged serial passage at high multiplicities (11, unpublished data) and sequences derived from host DNA are included with the SV40 *ori* regions of some of these variants. In addition, the efficiency of replication in these variants might be enhanced by duplication of small portions of the 21-bp repeat as well as an inversion of the 21-bp sequence. That is, if replication is efficiently initiated from the 111-bp *Hin* C segment of the passage 20 variant, it might depend on regulatory elements supplied in inverted orientation from the 260-bp *ori* segment. Although the 21-bp repeat is also not intact in the 119-bp segment of *Hin* C, the passage 40 variant does contain one copy of the consensus sequence and possibly part of another copy (due to a 4-bp duplication). In the passage 45 variant, one arm of the inverted repeat has a 6-bp GCG_2Pu_2 duplication (colinear with a portion of the consensus sequence) plus a 14-bp continuation of *Hin* C to complete the 133-bp *Hin* C segment for a total of two copies of the consensus sequence plus most (6 of 8 bp) of a third. The other arm of the passage 45 inverted repeat has four copies of the consensus sequence as a result of the 22-bp duplication of *Hin* C. The 69-bp host sequence present between the *ori* segments in *ev*-1104 and *ev*-1108 is not G+C rich and does not contain the $G_3CG_2Pu_2$ consensus sequence, although it does contain the 4-bp GCGG sequence (at 47 or 23 bp from either *ori* segment), which is also present as a duplicated sequence in the *ori* region of *ev*-1108. If indeed the 21-bp tandem repeats and multiple copies of a consensus sequence increase the replication efficiency, then it appears that through evolutionary change we are seeing the natural selection of a better regulatory region for replication. Furthermore, the inverted repetition of this region may also play an important role. We plan to test in what ways the occurrence of viral segments rearranged as inverted repeats, duplications and insertions, and host sequences interspersed with viral segments affect the replication efficiency of these variants.

By computer analysis, the host sequences in *ev*-2101 and *ev*-1101 share no homology with each other or with host sequences present in other SV40 variants characterized in our laboratory (12; unpublished data) and in other laboratories (7, 24, 30) or with highly repetitive α-component satellite monkey DNA (20) or the *Alu* family of dispersed repetitive sequences (14). Presumably, the host segments in *ev*-2101 and *ev*-1101 are derived from sequences that are less frequently repeated in the monkey genome.

Analyses of the viral-host recombinant joints in *ev*-2101 and *ev*-1101 demonstrate that although a few nucleotides of shared homology and A+T richness might increase the potential for formation of the recombinant molecule, they are not essential to the recombination event. Wilson et al. (32) recently reported that the intramolecular recombination events within an SV40-plasmid chimera in monkey cells were also minimally dependent on sequence homology. The more studies we (12, 33, 34) and others (2, 17, 25, 26) do the more clear it becomes that illegitimate recombination is the major mechanism by which SV40 recombines with the host cell DNA.

When looking for additional structural features of the sequence that might be involved in forming the initial recombinant intermediate, we find that there is a notable clustering of short repeats (direct, inverted, and palindromic) near the two viral-host recombinant junctions in *ev*-2101, whereas similar short repeats are more dispersed and further removed from the crossover of the two recombinant joints in *ev*-1101. The 300-bp cell insert in *ev*-1101 contains two sizeable (10 and 12 bp) sequence segments with twofold rotational symmetry. Within the 529-bp cell DNA insert in *ev*-2101 there are three palindromic sequences that are 10 bp or longer and an 11-bp sequence located at 102 bp from one viral-host junction that is inverted and repeated 24 bp from the other viral-host junction. Nordheim and Rich (19) have suggested that recombination may be facilitated by stretches of alternating purine and pyrimidine residues which favor a change in the conformation of the double helix from right-handed B-DNA to left-handed Z-DNA (31). Short segments (6 to 8 bp) of alternating purine and pyrimidine bases occurring in the cellular inserts and near the recombinant joints are noted in Fig. 2 to 4. Although it is interesting to point out these various sequence elements, it remains to be determined whether they are functionally significant.

ACKNOWLEDGMENTS

We thank Kenneth Manly for writing the computer program we used to search for sequence homologies and Marcia Held and Bonnie Beccue for typing the manuscript.

This research was supported by Public Health Service grant CA-28250 from the National Institutes of Health. M.W.-G. is the recipient of a Research Career Development Award (CA-00449) from the National Institutes of Health.

LITERATURE CITED

1. **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.
2. **Botchan, M., J. Stringer, T. Mitchison, and J. Sambrook.** 1980. Integration and excision of SV40 DNA from the chromosome of a transformed cell. Cell **20**:143–152.
3. **Brockman, W. W.** 1977. Evolutionary variants of simian virus 40. Prog. Med. Virol. **23**:69–95.
4. **Brockman, W. W., T. N. H. Lee, and D. Nathans.** 1973. The evolution of new species of viral DNA during serial passage of simian virus 40 at high multiplicity. Virology **54**:384–397.
5. **Brockman, W. W., and D. Nathans.** 1974. The isolation of simian virus 40 variants with specifically altered genomes. Proc. Natl. Acad. Sci. U.S.A. **71**:942–946.
6. **Danna, K. J., and D. Nathans.** 1972. Bidirectional replication of simian virus 40 DNA. Proc. Natl. Acad. Sci. U.S.A. **69**:3097–3100.
7. **Dhruva, B. R., T. Shenk, and K. N. Subramanian.** 1980. Integration in vivo into a simian virus 40 DNA of a sequence that resembles a certain family of genomic interspersed repeated sequences. Proc. Natl. Acad. Sci. U.S.A. **77**:4514–4518.
8. **DiMaio, D., and D. Nathans.** 1980. Cold sensitive regulatory mutants of SV40. J. Mol. Biol. **140**:129–142.
9. **Fareed, G. C., C. F. Garon, and N. P. Salzman.** 1972. Origin and direction of SV40 DNA replication. J. Virol. **10**:484–491.
10. **Gluzman, Y., J. Sambrook, and R. J. Frisque.** 1980. Expression of early genes of origin-defective mutants of SV40. Proc. Natl. Acad. Sci. U.S.A. **77**:3898–3902.
11. **Gutai, M. W., and D. Nathans.** 1978. Evolutionary variants of simian virus 40: nucleotide sequence of a conserved SV40 DNA segment containing the origin of viral DNA replication as an inverted repetition. J. Mol. Biol. **126**:259–274.
12. **Gutai, M. W., and D. Nathans.** 1978. Evolutionary variants of simian virus 40: cellular DNA sequences and sequences at recombinant joints of substituted variants. J. Mol. Biol. **126**:275–288.
13. **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.
14. **Jelinek, W. R., and C. W. Schmid.** 1982. Repetitive sequences in eukaryotic DNA and their expression. Annu. Rev. Biochem. **51**:813–844.
15. **Lee, T. N. H., W. W. Brockman, and D. Nathans.** 1975. Evolutionary variants of simian virus 40: cloned substituted variants containing multiple initiation sites for DNA replication. Virology **66**:53–69.
16. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavage. Methods Enzymol. **65**:499–560.
17. **McCutchan, T., M. Singer, and M. Rosenberg.** 1979. Structure of simian sirus 40 recombinants that contain both host and viral DNA sequences. II. The structure of variant 1103 and its comparison to variant CVP8/1/P2 (*Eco*RI res). J. Biol. Chem. **254**:3592–3597.
18. **Myers, R. M., and R. Tjian.** 1980. Construction and analysis of SV40 origins defective in tumor antigen binding and DNA replication. Proc. Natl. Acad. Sci. U.S.A. **77**:6491–6495.
19. **Nordheim, A., and A. Rich.** 1983. The sequence $(dC\text{-}dA)_n$ $(dG\text{-}dT)_n$ forms left-handed Z-DNA in negatively supercoiled plasmids. Proc. Natl. Acad. Sci. U.S.A. **80**:1821–1825.
20. **Rosenberg, H., M. F. Singer, and M. Rosenberg.** 1978. Highly reiterated sequences of SIMIANSIMIANSIMIANSIMIANSIMIAN. Science **200**:394–402.
21. **Sanger, F., and A. R. Coulson.** 1978. The use of thin acrylamide gels for DNA sequencing. FEBS Lett. **87**:107–110.
22. **Shortle, D., and D. Nathans.** 1979. Regulatory mutants of SV40: constructed mutants with base substitutions at the origin of DNA replication. J. Mol. Biol. **131**:801–817.
23. **Sol, C. J. A., I. Hassing, W. Maris, C. Walig, and J. van der Noordaa.** 1981. Evolutionary variants of simian virus 40 which are impaired in early lytic functions but transform nonpermissive cells. J. Virol. **37**:395–410.
24. **Sompayrac, L., and K. J. Danna.** 1983. Simian virus 40 deletion mutants that transform with reduced efficiency. Mol. Cell. Biol. **3**:484–489.
25. **Stringer, J. R.** 1981. Integrated simian virus 40 DNA: nucleotide sequences at cell-virus recombinant junctions. J. Virol. **38**:671–679.
26. **Stringer, J. R.** 1981. DNA sequence homology and chromosomal deletion at a site of SV40 DNA integration. Nature (London) **296**:363–366.
27. **Subramanian, K. N., and T. Shenk.** 1978. Definition of the boundaries of the origin of DNA replication in SV40. Nucleic Acids Res. **5**:3635–3642.
28. **Tjian, R.** 1979. Protein-DNA interactions at the origin of SV40 DNA replication. Cold Spring Harbor Symp. Quant. Biol. **43**:655–662.
29. **Tooze, J. (ed).** 1980. DNA tumor viruses: the molecular biology of tumor viruses, part 2. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
30. **Wakamiya, T., T. McCutchan, M. Rosenberg, and M. F. Singer.** 1979. Structure of simian virus 40 recombinants that contain both host and viral DNA sequences. I. The structure of variant CVP8/1/P2 (*Eco*RI res). J. Biol. Chem. **254**:3584–3591.
31. **Wang, A. H.-J., G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel, and A. Rich.** 1979. Molecular structure of left-handed double helical DNA fragment at atomic resolution. Nature (London) **282**:680–687.
32. **Wilson, J. H., P. B. Berget, and J. M. Pipas.** 1982. Somatic cells efficiently join unrelated DNA segments end-to-end. Mol. Cell. Biol. **2**:1258–1269.
33. **Woodworth-Gutai, M.** 1981. Recombination in SV40-infected cells: nucleotide sequences at viral-viral recombinant joints in naturally arising variants. Virology **109**:344–352.
34. **Woodworth-Gutai, M.** 1981. Recombination in SV40-infected cells: viral DNA sequences at sites of circularization of transfecting linear DNA. Virology **109**:353–365.