MOLECULAR AND CELLULAR BIOLOGY, June 1985, p. 1247–1259
0270-7306/85/061247-13$02.00/0

Vol. 5, No. 6

Recombination Between Poly[d(GT) · d(CA)] Sequences in Simian Virus 40-Infected Cultured Cells

JAMES R. STRINGER

Department of Microbiology and Molecular Genetics, University of Cincinnati College of Medicine, Cincinnati, Ohio 45267

Received 26 December 1984/Accepted 7 March 1985

CVI cells were transfected with oversized simian virus 40 (SV40) genomes that could be reduced to packageable size by alternative homologous recombination pathways involving either two polydeoxyguanylic-thymidylic acid · polydeoxycytidylic-adenylic acid {poly[d(GT) · d(CA)]; abbreviated hereafter as poly(GT)} tracts or two tracts of homologous SV40 sequence. Plaque-forming viruses rescued by this procedure were found to contain genomes formed by homologous and nonhomologous recombination events. Half of the viable viral DNA molecules recovered were the result of recombination between two tracts of poly(GT). Approximately 20% of the rescued viral genomes were produced by homologous recombination between tracts of SV40 DNA. Nonhomologous recombination involving SV40 sequences was also a major pathway of deletion, producing ca. 30% of the viral plaques. Tracts of poly(GT) generated by recombination were variable in length, suggesting that recombination between poly(GT) tracts was usually unequal. On a per-nucleotide basis, poly(GT) recombination occurred eight times more frequently than did recombination between homologous SV40 DNA. This eightfold difference is the maximum recombinatory enhancement attributable to poly(GT) sequences. Although DNA sequence analysis showed that tracts of poly(GT) generated by recombination retained the alternating G-T repeat motif throughout their length, the contribution of the nonhomologous pathway to poly(GT) recombination cannot be ruled out, and the relative proclivity of a given length of d(GT) · d(CA) sequence to undergo homologous recombination is probably less than eight times greater than that of an SV40 sequence of the same length.

Eucaryotic genomes contain numerous runs of alternating deoxyguanosine and thymidine residues {polydeoxyguanylic-thymidylic acid · polydeoxycytidylic-adenylic acid, poly [d(GT) · d(CA)]; abbreviated hereafter as poly(GT)} 10 to 50 base pairs in length (9–11). In mammals, poly(GT) sequences are dispersed throughout the genome and have been estimated to number on the order of 10^5 copies per cell (9, 10). Many cloned genes have been found to contain poly(GT) either in introns or in other untranslated regions, and additional examples of poly(GT) have been found near a number of other cloned genes (7, 16, 18, 22, 26, 27, 29, 31, 33, 34, 39). Poly(GT) sequences adopt a left-handed Z conformation when placed under sufficient torsional strain by supercoiling in vitro (14, 28). Although the existence of Z DNA in vivo remains at issue, the behavior of poly(GT) sequences may be indicative of biological activities associated with sequences that exhibit a propensity to isomerize in vitro. Poly(GT) sequences have been shown to weakly enhance transcription in transient expression experiments (12). An altogether different role for poly(GT) sequences has been suggested based on sequence analysis of human γ globin genes which show evidence of having undergone gene conversion events in a region containing a poly(GT) sequence (33). Additional circumstantial evidence linking poly(GT) with recombination has been previously reported (7, 33, 35, 38).

The experiments described in this report examined the propensity of poly(GT) sequences to undergo recombination in cultured monkey cells. CVI cells were transfected with oversized, poly(GT)-containing simian virus 40 (SV40) genomes that could be reduced to packageable size by alternative homologous recombination pathways involving either two poly(GT) tracts or two homologous SV40 sequences. Rescued viruses were found to have been formed by deletions in the input DNA that were mediated by homologous and nonhomologous recombination events. Half the viable viral DNA molecules recovered were the result of recombination between two tracts of poly(GT). Approximately 20% of the rescued viral genomes were produced by homologous recombination between tracts of SV40 DNA. Nonhomologous recombination involving SV40 sequences was also a major pathway of deletion, producing ca. 30% of the viral plaques. This illegitimate deletion pathway was not due to recombination between SV40 and poly(GT) sequences. The products of recombination between two poly(GT) tracts were variable in length and retained the dinucleotide repeat pattern, suggesting that homologous crossing-over between misaligned base-paired recombination intermediates was primarily responsible for poly(GT) recombination. However, breakage and reunion of two poly(GT) tracts mediated by the nonhomologous pathway would be cryptic half the time, and some nonhomologous recombination of poly(GT) tracts may have gone undetected. Although the prevalence of nonhomologous recombination created uncertainty in the assessment of the relative propensity of poly(GT) sequences to recombine homologously, the data allow estimation of the maximum ratio of recombinatory potential for poly(GT) versus SV40 DNA. Poly(GT) recombination occurred eight times more frequently on a per-nucleotide basis than did homologous recombination between two tracts of SV40 DNA. This suggests that a given length of poly(GT) sequences is no more than eight times more likely to undergo homologous recombination than an SV40 DNA sequence of equivalent length.

MATERIALS AND METHODS

Enzymes. All enzymes were used as described by Maniatis et al. (23). Commercial sources of enzymes were as follows. Restriction endonucleases *Cla*I, *Xba*I, *Eco*RI, *Hin*dIII, *Hpa*I, *Pst*I, *Mbo*I, *Sau*3A, *Alu*I, and *Bcl*I, T4 DNA ligase, and nuclease Bal 31 were from New England Biolabs (Beverly, Mass.). *Escherichia coli* DNA polymerase I and *E. coli* DNA polymerase I Klenow fragment were from Bethesda Research Laboratories (Gaithersburg, Md.).

DNA. DNA linkers containing *Cla*I and *Xba*I recognition sequences were purchased from Collaborative Research, Inc. (Lexington, Mass.) and used as described by Maniatis et al. (23). Plasmid DNA was prepared by the boiling lysis procedure, followed by equilibrium centrifugation in CsCl in the presence of ethidium bromide (23). DNA fragments produced by restriction endonuclease cleavage were separated by size via electrophoresis through gels composed of agarose or polyacrylamide as previously described (23). Purification of DNA fragments from gels was by electroelution and anion-exchange chromatography as previously described (23). DNA fragment concentrations were estimated by comparative visual assessment of fluorescence of bound eithidium bromide (23). Ligation of DNA was performed at 16°C in a buffer composed of 50 mM Tris-hydrochloride (pH 7.8), 10 mM $MgCl_2$, 20 mM dithiothreitol, and 1.0 mM ATP at a total DNA concentration of ca. 50 μg/ml. Multifragment ligation reactions contained approximately equimolar quantities of each DNA fragment. Ligation of recombination substrates and control SV40 DNA contained SV40 *Pst*I A fragment from the same stock solution. The SV40 *Pst*I B fragment and modified versions of the SV40 *Pst*I B fragment were used in equivalent amounts. The extent of ligation was monitored by end labeling with deoxynucleotide triphosphates labeled with ^{32}P at the α phosphate (New England Nuclear Corp., Boston, Mass.), followed by agarose gel electrophoresis and autoradiography as described previously (23).

SV40 nucleotide numbering system. SV40 nucleotides are numbered as described in the system in reference 3.

Bacterial transformation. Transformation of bacteria was by the protocol described by Hanahan (13). Recipient strains were DH-1 (F^- *recA1 endA1 gyrA96 thi-1 hsdR17* r_K^- m_K^+ *supE44 relA*? λ^-) and W3110 (F^- *hsdR hsdM*$^+$).

DNA constructions. $SV(GT)^2$ was prepared as follows. A plasmid, pHA, which contained the *Hin*dIII A fragment of SV40 cloned into the *Hin*dIII site of the plasmid pAT, was modified to produce three derivatives, pHAX, pHAC, and pHAGT. To make pHAX and pHAC, the unique *Hpa*I cleavage site in pHA was changed to either an *Xba*I site or a *Cla* site by ligation of linkers containing either of these cleavage sites. pHAGT was produced when a 135-base-pair DNA fragment containing poly(GT) was ligated into the *Hpa*I site of pHA. Plasmid pXGTC was produced by insertion of an 85-base-pair poly(GT) segment into the *Cla* site of pBR322, followed by cleavage with *Eco*RI, treatment with Bal 31, and ligation of an *Xba* linker. A plasmid, $p(GT)^2$ (see Fig. 2), was constructed by ligation of a mixture containing the large *Eco*RI-*Hin*dIII fragment of pBR322, the small *Eco*RI-to-*Bcl* fragment of pHAGT, the small *Bam-Xba* fragment of pHAX, the small *Xba-Cla* fragment of pXGTC, and the *Cla-Hin*dIII fragment of pHAC, corresponding to SV40 sequences from the *Hpa*I site at SV40 nucleotide 2666 to the *Hin*dIII site at SV40 nucleotide 3459. $SV(GT)^2$ (see Fig. 1) was made by ligation of the small *Pst*I fragment of $p(GT)^2$ to the large *Pst*I fragment of SV40 DNA. $SV(AB)^3$ (see Fig. 3) was prepared as follows. A plasmid designated $p(AB)^2$ was made by ligation of the large *Eco*RI-*Hin*dIII fragment of pBR322, the small *Eco*RI-*Bcl* fragment of SV40, and the 943-base-pair *Bam*HI-*Hin*dIII fragment of SV40. $p(AB)^2$ was converted to $p(AB)^3$ by ligation of the large *Eco*RI-*Bam* fragment of $p(AB)^2$ to the small *Eco*RI-*Bcl* fragment of SV40. $SV(AB)^3$ DNA was produced by ligation of the small *Pst*I fragment of $p(AB)^3$ with the large *Pst*I fragment of SV40.

Transfection of CVI cells. CVI cells were transfected by the DEAE-dextran method (21).

Plaque assay. Plaques were formed on monolayers of CVI cells that had been overlaid with 0.9% agarose 24 h postinfection. Visualization of plaques was aided by neutral red staining.

Viral DNA preparation. Viral DNA was prepared from CVI cells infected with virus derived from agarose plugs pulled from the area overlaying viral plaques produced after transfection of CVI cells with $SV(GT)^2$ or similar DNA. Infected cultures were harvested when viral cytopathic effect was apparent, and they were processed by the method of Hirt (15).

DNA analysis. Restriction endonuclease mapping was as described by Maniatis et al. (23). DNA fragments were separated by electrophoresis through agarose (Sigma Chemical Co., St. Louis, Mo.) or polyacrylamide (Bio-Rad Laboratories, Richmond, Calif.) in TBE buffer (0.09 M Tris, 0.09 M boric acid, 0.005 M EDTA [pH 8]). DNA sequencing was as described by Maxam and Gilbert (25). DNA samples to be sequenced were end labeled by treatment with the Klenow fragment of *E. coli* DNA polymerase I. Class A molecules were end labeled at the *Xba* site. Class B molecules were end labeled at the *Bcl*I site. Class GT molecules were end labeled at the *Cla* site. Sequence reactions were G, G+A, C+T, and C. The G+A modification reaction used 50% formic acid.

Statistical methods. The statistical significance of variations from expected frequencies was evaluated by the Student *t* test as described previously (5).

RESULTS

Isolation of recombinants by virus rescue. Previous work from this laboratory had shown that poly(GT) could be inserted into the SV40 genome without the loss of viral viability (36). Briefly, a 135-base-pair insertion, of which 90 base pairs were poly(GT), was inserted at nucleotide 2666 of the SV40 genome to produce a virus termed SVGT. The poly(GT) insertion was mildly deleterious, causing SVGT to grow more slowly than SV40, and it was therefore somewhat surprising to find that SVGT could be propagated without a detectable loss of poly(GT). Such a loss might be expected to occur by homologous recombination within the internally repetitious run of GT dinucleotides. Alternative explanations for the stability of poly(GT) in SVGT are possible. One possibility is that recombination within the GT repeat is rare or absent. The alternative explanation is that deletions do occur, but that such deletions do not relieve the deleterious effect of the insertion to a degree sufficient to allow deleted genomes to outgrow SVGT. The experiment diagrammed in Fig. 1 was designed to distinguish between these alternatives by measurement of the recombination frequencies of tandem poly(GT) tracts relative to the recombination frequency between tandemly arrayed tracts of homologous SV40 DNA.

Because SV40 virions do not readily accommodate much more than 5,500 base pairs of DNA (1, 2, 41), it was

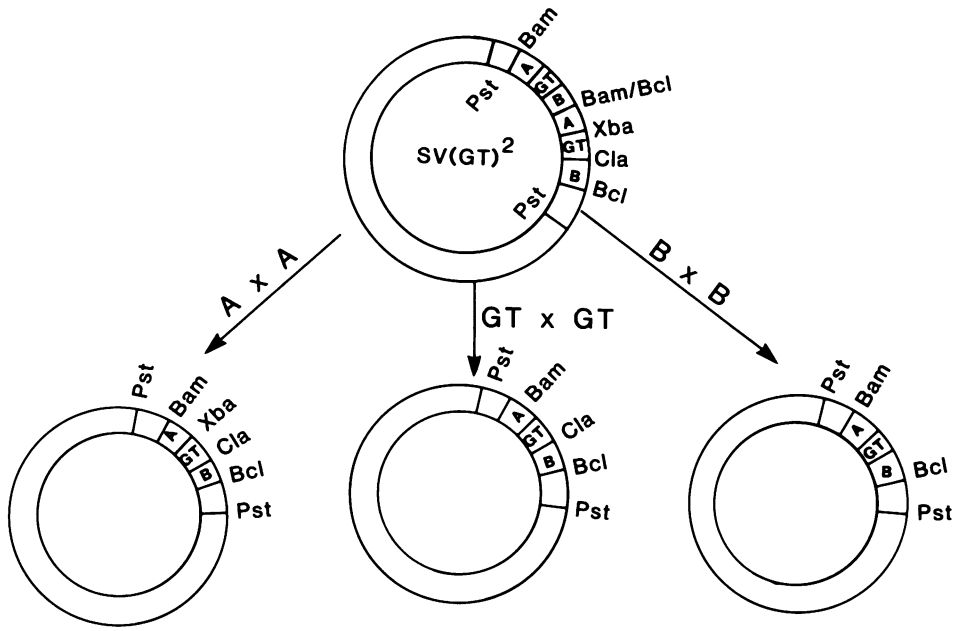


FIG. 1. Structures of the recombination substrate SV(GT)^2 and the molecules deleted via alternative homologous recombination pathways. *Bam*, *Xba*, *Cla*, *Bcl*, and *Pst* indicate the location of cleavage sites for the restriction endonucleases *Bam*HI, *Xba*I, *Cla*I, *Bcl*I, and *Pst*I, respectively. Intervals marked A represent the 133 base pairs of SV40 sequence that lie between nucleotides 2533 and 2666. Intervals marked B represent the 104 base pairs between nucleotides 2666 and 2770 of the SV40 genome. Intervals marked GT contained 85 to 90 base pairs of poly(GT) DNA. Arrows labeled A×A, GT×GT, and B×B point to the structures that would be generated by homologous recombination between two A segments, two poly(GT) segments, or two B segments.

reasonable to expect that the packaging limit of SV40 virion formation could be used to select for recombination events that would delete DNA from an oversized viral genome. The DNA molecule diagrammed at the top of Fig. 1 [SV(GT)^2] was constructed to serve as an unpackageable viral genome that could delete, by homologous recombination, one copy of a tandemly repeated segment of the viral genome. The repeated segment contained the SV40 sequences that lie between the cleavage sites for *Bam*HI and *Bcl*I. Embedded at nucleotide 2666 in each of the two copies of the SV40 *Bam-Bcl* fragment was a poly(GT) insertion. One poly(GT) insertion was flanked by cleavage sites for the restriction endonucleases *Xba*I and *Cla*I and contained 85 base pairs of poly(GT). The other poly(GT) insertion lacked restriction endonuclease cleavage sites at its ends and contained 90 base pairs of poly(GT) and 45 base pairs of DNA composed of a more complex sequence. The sequence of these 45 base pairs has been reported previously (36). SV(GT)^2 DNA is structurally disposed to delete by homologous recombination to produce packageable SV40 genomes with a structure similar to the viable variant, SVGT. Theoretically, deletion could occur either by recombination between repeated poly(GT) tracts or by recombination between either of the pairs of homologous SV40 sequences flanking the poly(GT) insertion (labeled A and B in Fig. 1). Which sequences participated in a recombination event should be indicated by the segregation of *Cla*I and *Xba*I cleavage sites. The presence of both a *Cla*I cleavage site and an *Xba*I cleavage site would indicate that deletion had occurred by recombination between the SV40 sequences designated A. The presence of a *Cla*I site and the absence of an *Xba*I site would indicate recombination between poly(GT) tracts. The absence of both *Cla*I and *Xba*I cleavage sites would be expected in progeny molecules produced by recombination between the SV40 sequences labeled B.

The protocol for construction of the oversized SV40 genome, its introduction into cells, and selection of recombinant plaques is diagrammed in Fig. 2. Transfection of CVI cells with the SV(GT)^2 construction yielded ca. 400 plaques per μg of SV(GT)^2 DNA. This yield was 2.5-fold lower than that achieved by transfection with SV40 genomes that had been reconstructed by ligation of gel-purified DNA fragments produced by *Pst*I digestion of SV40 DNA. The diminution in yield of plaques from SV(GT)^2 was correlated with the requirement that the input DNA delete and not with the presence of poly(GT) in the recombination substrate. This was shown by transfection of CVI cells with an oversized viral genome, SV(AB)^3 (Fig. 3), that was the same size as SV(GT)^2 but lacked poly(GT) insertions. The efficiency of rescue of virus plaques from SV(AB)^3 DNA was the same as that found for SV(GT)^2 DNA. Plaques rescued from SV(GT)^2 and SV(AB)^3 DNAs were slower to form than those rescued from ligated genomes that need not delete to package. The delay associated with virus rescue from recombination substrates was ca. 2 days.

Rescued viruses produced by recombination in CVI cells, not in *E. coli*. To assess the possible contribution of deletion in *E. coli* to the rescue of virus, the cloned DNA fragment used in the construction of SV(GT)^2 DNA (Fig. 2) was analyzed for purity. DNA was digested with *Bam*HI, radioactively end labeled, and electrophoresed through 2.5% agarose. The gel was dried onto filter paper, and the position of DNA fragments was detected by autoradiography. Under conditions in which a contamination of 1 part per 500 would have been visible, no deleted fragments were detectable in the DNA used in the construction of SV(GT)^2 DNA (data not shown). Since the efficiency at which virus was rescued from SV(GT)^2 DNA was only 2.5-fold lower than the efficiency at which plaques were produced by ligated SV40 DNA, DNA deleted in *E. coli* could not have produced more than 1

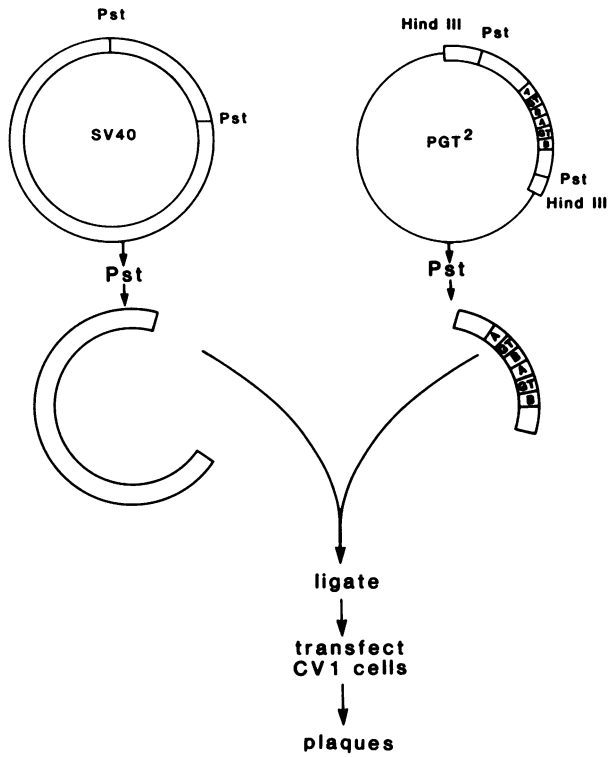


FIG. 2. Protocol for the rescue of virus plaques from oversized viral genomes. *Pst* and *Hin*dIII indicate the cleavages sites for the restriction endonucleases *Pst*I and *Hin*dIII, respectively. pGT^2 is a plasmid DNA bearing an SV40 *Hin*dIII A fragment that was modified to contain tandem repeats of SV40 and poly(GT) DNA. Segments labeled A, GT, and B are as described in the legend to Fig. 1. The DNA fragments used in the ligation step were purified by electrophoresis through agarose.

plaque for every 200 plaques formed by recombination in CVI cells. The delay in the appearance of plaques formed by recombination is consistent with the conclusion that recombination occurred in CVI cells and not in *E. coli*.

Rescued viruses produced by a variety of recombination events. Twenty-two $SV(GT)^2$-derived plaques were picked, and viral DNA was produced by infecting fresh cultures of CVI cells. Viral DNA was prepared by Hirt extraction and analyzed by digestion with *Cla*I and *Xba*I (Fig. 4). To eliminate false-negatives in the restriction endonuclease assay, samples found to be resistant to *Cla*I and *Xba*I cleavage were tested again in a mixture that contained a marker DNA that contained cleavage sites for *Cla*I and *Xba*I. In every case, the marker molecule was cleaved, and the viral DNAs were again resistant to digestion by *Cla*I and *Xba*I (data not shown).

The data from Fig. 4 are tabulated in Table 1. These data indicated that more viruses were the product of recombination in the poly(GT) sequence lying between *Cla*I and *Xba*I in $SV(GT)^2$ than would be expected if target size were the only determinant of rescue frequency. Conversely, relative to expectations based on target size, fewer viruses were rescued via recombination involving the A and B sequence intervals linked to the *Xba*I and *Cla*I sites in $SV(GT)^2$. In considering these data, it is important to point out that although the segregation of the diagnostic restriction sites, *Cla*I and *Xba*I, unequivocally indicates the location of one of the sequences that participated in a recombination event

TABLE 1. Analysis of viral genomes rescued from $SV(GT)^2$ DNA

Sequence interval	Length in base pairs	No. of plaques		Significance[a]
		Expected	Observed	
A	133	9.1	5	$P < 0.1$
GT	85	5.8	12	$P < 0.01$
B	104	7.1	5	$P < 0.4$

[a] Significance was calculated by the formula $Z = P - \theta \times [\theta(1 - \theta)/n]^{-1}$ (5) in which P is the observed proportion of plaques in a class, θ is the theoretical proportion of plaques in a population formed by homologous recombination at rates directly proportional to the relative sequence interval, and n is the sample size. Probabilities were adjusted for sample size effects by the Student t test formula (5).

resulting in a deletion, whether or not the other deletion endpoint can also be mapped on the basis of sensitivity to *Cla*I or *Xba*I depends upon the nature of the recombination event that mediated the deletion. If deletion were to occur by homologous recombination between tandemly arrayed blocks of identical sequences, the determination of one deletion endpoint would implicitly map the other endpoint to the DNA segment containing the homologous sequence. However, were deletion to occur by nonhomologous recombination, sensitivity to *Cla*I, *Xba*I, or both would map only the deletion endpoint linked to the diagnostic restriction sites. As will be demonstrated below, several of the viruses rescued from $SV(GT)^2$ DNA were the product of nonhomologous recombination. The prevalence of nonhomologous recombination created ambiguity in the relationship of virus rescue to recombination frequency because not all nonhomologous deletions would be expected to produce plaques. Therefore, quantitation of the relative propensities of poly(GT) and SV40 sequences to recombine homologously required further analysis of rescued viral genomes to distinguish homologous from nonhomologous events.

Prevalence of nonhomologous recombination in the deletion of oversized SV40 DNA. Figure 5 shows the results of

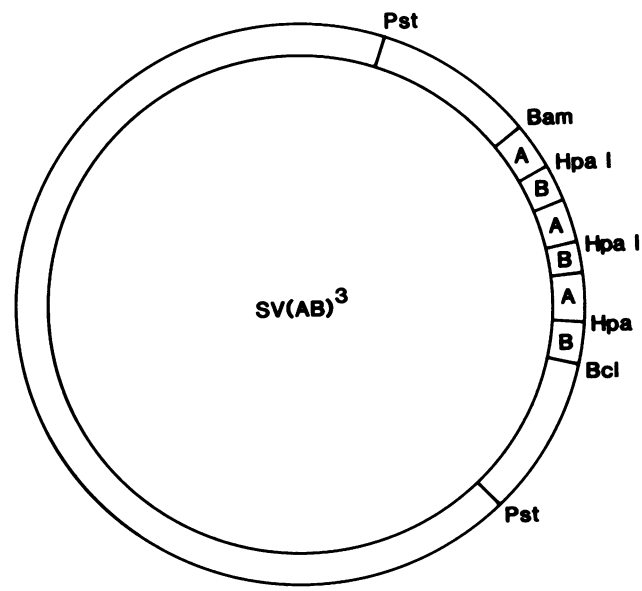


FIG. 3. The structure of recombination substrate $SV(AB)^3$. Segments designated A and B are as described in the legend to Fig. 1.

digestion of 13 rescued genomes with *Pst*I, which produced two fragments, the smaller of which contained the region of deletion. Deletion mediated by homologous recombination between two A segments or two B segments would produce *Pst*I fragments of discrete size: 1,301 base pairs in the case of A-by-A recombination and 1,351 base pairs in the case of B-by-B recombination. Of the five recombinants possessing both *Cla*I and *Xba*I sites (A-class), only two (A1 and A10b) produced *Pst*I fragments of the expected size. The other three A-class recombinants contained *Pst*I fragments larger than would be expected to be produced by homologous recombination between two A segments. Two examples of oversized A-class genomes are shown in lanes 7 and 13 of Fig. 5. Similarly, of the five recombinants lacking both the *Cla*I and *Xba*I cleavage sites (B-class), only two (B6c and B8e) produced *Pst*I fragments of the size expected to result from homologous recombination between two B segments.

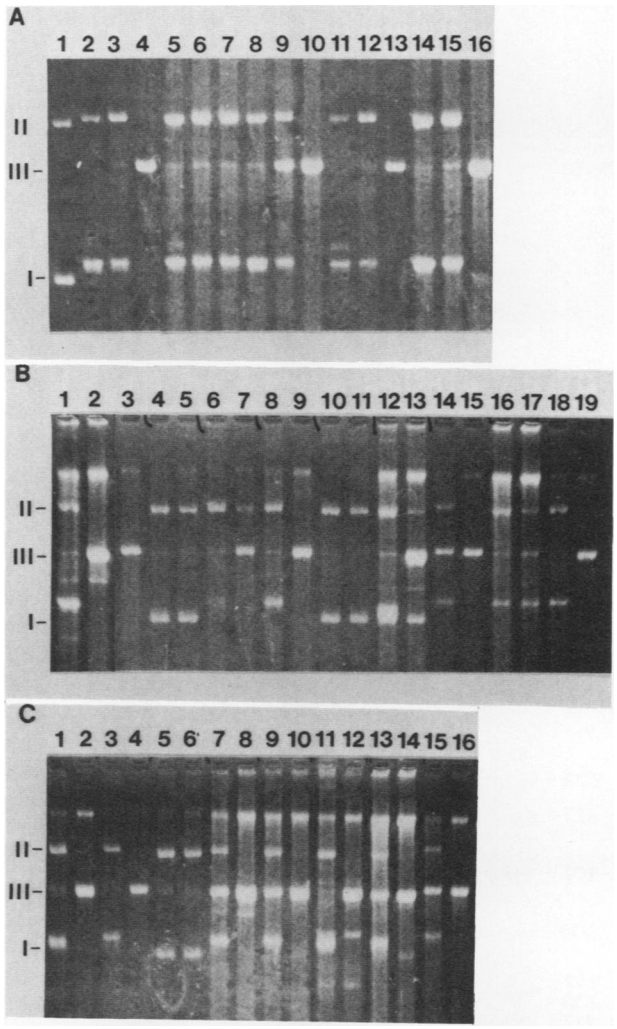


FIG. 4. *Cla* and *Xba* analysis of viral genomes rescued from SV(GT)^2 DNA. Viral DNA was prepared by Hirt extraction. DNA was electrophoresed through 0.7% agarose in TBE. DNA was visualized by staining with ethidium bromide. Roman numerals mark the position of supercoiled circular (I), relaxed circular (II), and linear (III) DNAs. (A) Lane 1, SV40 DNA; lanes 2, 5, 8, 11, and 14, undigested samples of isolates GT8a, B8e, A9c, GT9d, and GT10a; lanes 3, 6, 9, 12, and 15, same samples digested with the *Xba*I; and lanes 4, 7, 10, 13, and 16, same samples digested with *Cla*I. (B) Lanes 1, 4, 6, 8, 10, 12, 14, 16, and 18, *Xba*I-digested samples of isolates GT1a, B3c, GT3b, GT4b, B5b, GT6a, GT8b, B6c, and GT6d; lanes 2, 5, 7, 9, 11, 13, 15, 17, and 19, same samples digested with *Cla*I. (C) Lanes 1, 3, 5, 7, 9, 11, 13, and 15, *Xba*I-digested samples of isolates GT7b, GT9a, B9b, A10b, A1, GT3, A5, and A6b; lanes 2, 4, 6, 8, 10, 12, 14, and 16, same samples digested with *Cla*I.

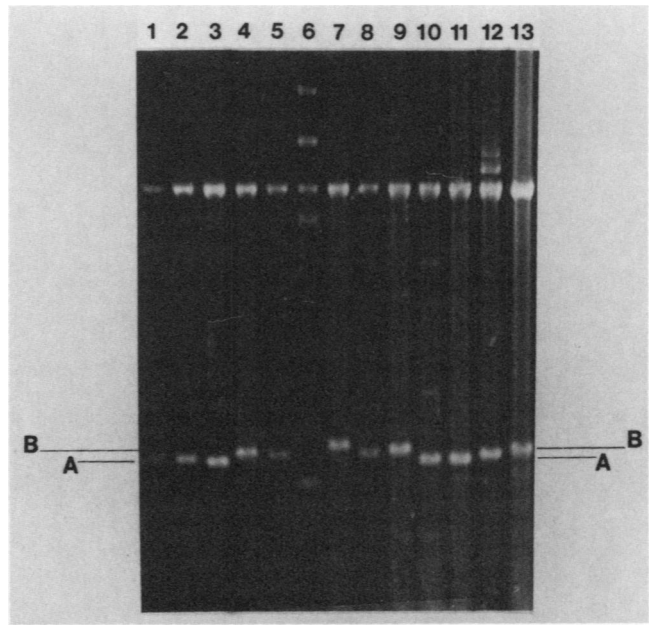


FIG. 5. *Pst*I analysis of viral genomes rescued from SV(GT)^2 DNA. DNA samples were prepared by Hirt extraction of infected cells. Electrophoresis was through 1% agarose in TBE. DNA was visualized with ethidium bromide. A and B mark the migration positions of the small *Pst*I fragments of viral genomes generated by homologous recombination between either two A segments or two B segments in SV(GT)^2 DNA. Lanes 1 through 13 contained samples GT7b, GT8a, GT8b, B8e, GT9a, B9b, A9c, GT9d, GT10a, A10b, A1, GT3, and A5, respectively.

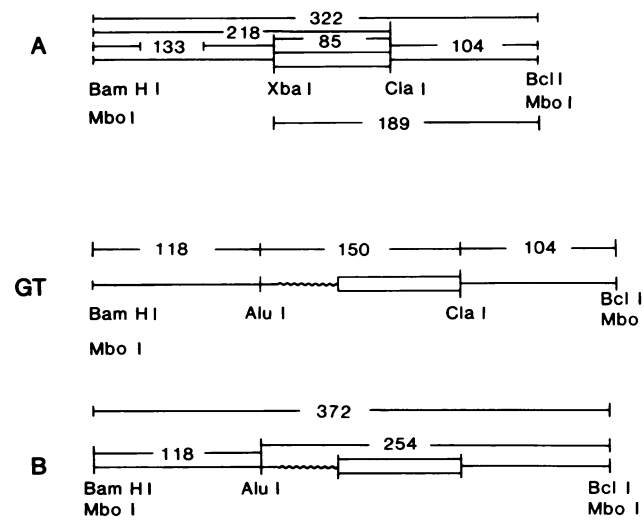
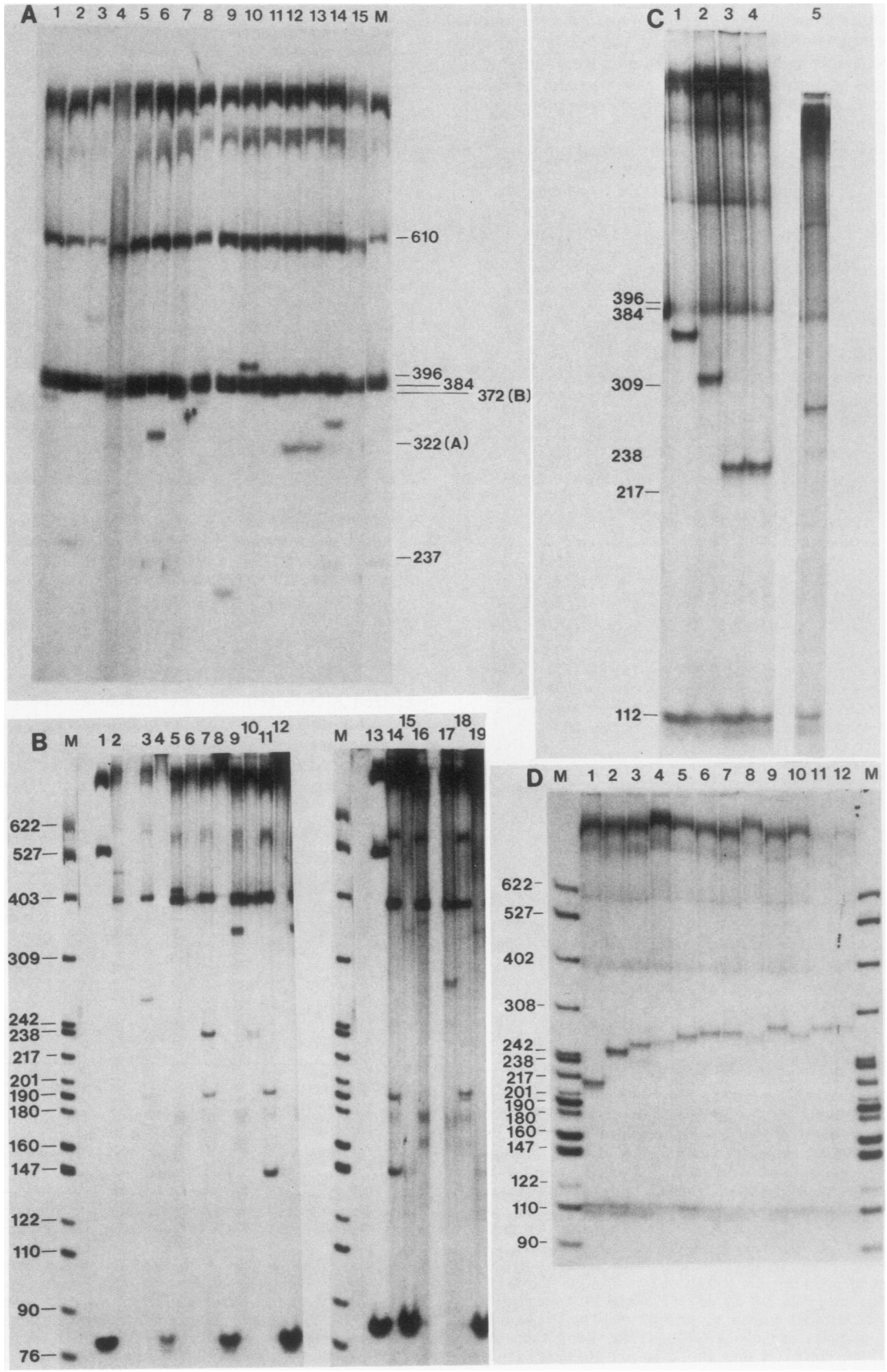


FIG. 6. Restriction maps of the *Bam-Bcl* fragments expected to be generated by homologous recombination events. The 150-base-pair length between the *Alu* and *Cla* sites is what would be expected if GT-by-GT recombination was equal. Unequal exchanges between GT tracts would produce variation in the length of *Alu-Cla* fragments. Symbols: (▭) poly(GT); (—) SV40; and (〰) non-poly(GT) insertion.


A
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 M
610
396
384
372 (B)
322 (A)
237
C
1 2 3 4 5
396
384
309
238
217
112
B
M 1 2 3 4 5 6 7 8 9 10 11 12
M 13 14 15 16 17 18 19
622
527
403
309
242
238
217
201
190
180
160
147
122
110
90
76
D
M 1 2 3 4 5 6 7 8 9 10 11 12 M
622
527
402
308
242
238
217
201
190
180
160
147
122
110
90

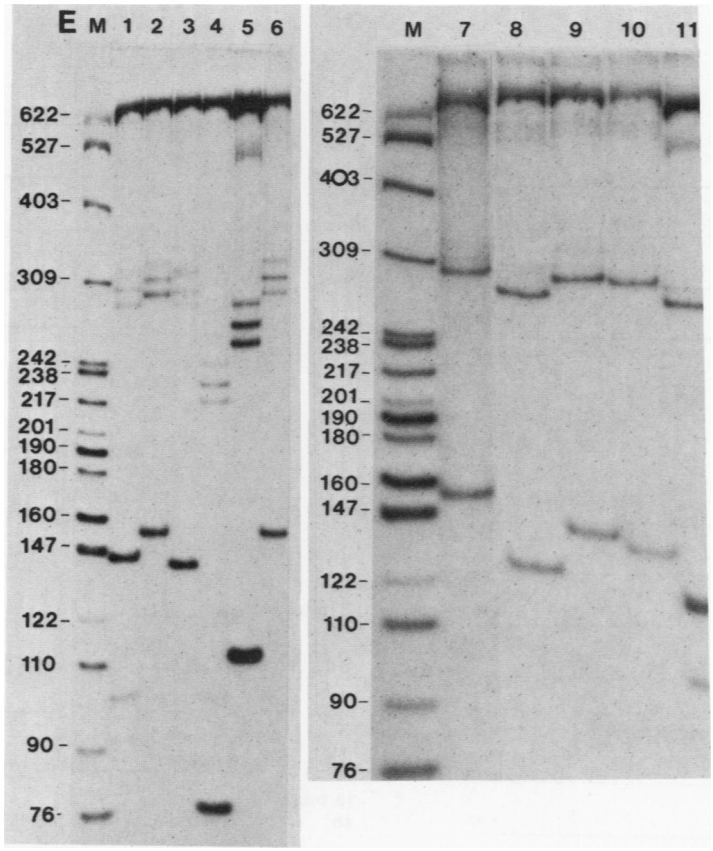


FIG. 7. Restriction endonuclease mapping of viral genomes rescued from SV(GT)^2 DNA. All panels are autoradiograms of end-labeled DNA fragments separated by electrophoresis through polyacrylamide. The numbers in the margins indicate the approximate molecular sizes of DNA fragments (in base pairs). Unless otherwise indicated, lanes marked (M) contained end-labeled *Hpa*II fragments of pBR322 DNA. The doublet bands seen in some of the smaller fragments are an artifact of the end-labeling procedure. (A) *Mbo*I digestion. Lanes 1 through 15 contained DNA from isolates GT4b, B5b, A6b, B6c, GT6d, GT8b, B8e, GT9a, B9b, A9c, GT10a, A1, A10b, GT3, and A5, respectively. Lane M contained SV40 DNA. (B) Double digestion of A-class molecules with either *Mbo*I and *Xba*I or *Xba*I and *Cla*I. Lanes 3, 7, 11, 14, and 18 contained samples of isolates A6b, A9c, A10b, A1, and A5 that had been digested first with *Mbo*I, end labeled, and subsequently cut with *Xba*I. Lanes 4, 8, 12, 15, and 19 contained samples of isolates A6b, A9c, A10b, A1, and A5 that had been digested first with *Xba*I, end labeled, and subsequently cut with *Cla*I. Lanes 1 and 13 contained the 85-base-pair poly(GT) sequence bounded by *Cla* and *Xba* in SV(GT)^2 DNA (see Fig. 1). (C and D) Samples were end labeled at the *Cla*I site and subsequently digested with *Mbo*I. (C) A-class molecules. Lanes 1 through 5 contained samples A6b, A9c, A10, A1, and A5, respectively. (D) GT-class molecules. Lanes 1 through 12 contained samples GT6a, GT8b, GT3, GT8a, GT7b, GT4b, GT1a, GT9d, GT10a, GT9a, GT3b, and GT6d, respectively. (E) GT-class samples were end labeled at the *Cla*I site and subsequently digested with *Alu*I. Lanes 1 through 11 contained samples GT1a, GT3b, GT4b, GT6a, GT8b, GT10a, GT6d, GT8a, GT9a, GT9d, and GT3, respectively. The families of bands migrating at ca. 300 base pairs are the result of partial *Alu*I digestion.

The other three B-class molecules contained undersized *Pst*I fragments. An example of a B-class viral genome with a small *Pst* fragment is shown in lane 6 of Fig. 5. These data indicated that most of the molecules assigned to either the A or B class (Table 1) were not the product of homologous recombination.

Mapping with restriction endonucleases that cut more proximally to the points of recombination allowed further localization of recombination points (see Fig. 6 for relevant restriction maps). Figure 7A shows the sizing of fragments produced by *Mbo*I cleavage of recombinant viral genomes. The data in Fig. 7A show that the size heterogeneity of rescued viral genomes of all classes was confined to the sequences between the *Bam* and *Bcl* sites. Consistent with the data derived by *Pst*I digestion, only two of the class A molecules (lanes 12 and 13) and two of the class B molecules (lanes 4 and 7) contained *Mbo*I fragments of the size expected to be produced by homologous recombination. The other A-class molecules contained *Mbo*I fragments larger than would be expected to result from homologous A-by-A recombination (lanes 3, 10, and 15; the *Mbo* fragment in lane 15 comigrated with the 384-base-pair SV40 fragment), and the other B-class viruses contained *Mbo*I fragments smaller than would be expected to result from homologous B-by-B recombination. Lanes 2 and 9 of Fig. 7A show two examples of such B-class viruses. Class B recombinants were further characterized by digestion with *Bam* plus *Alu* and *Bcl* plus *Alu* (data not shown). Cleavage with these enzymes showed the size heterogeneity in B-class genomes to lie between the *Alu* and *Bcl* sites. Three class B recombinants were further analyzed by DNA sequencing (see below and Fig. 10).

The results of further analysis of A-class and GT-class recombinant genomes are shown in Fig. 7B through E. The data in Fig. 7B and C mapped the size heterogeneity of class A molecules to the sequences bounded by *Bam* and *Xba* sites. Class A molecules all contained the expected 85-base-pair *Xba-Cla* fragment (Fig. 7B, lanes 4, 8, 12, 15, and 19), the expected 189-base-pair fragment bounded by *Xba* and

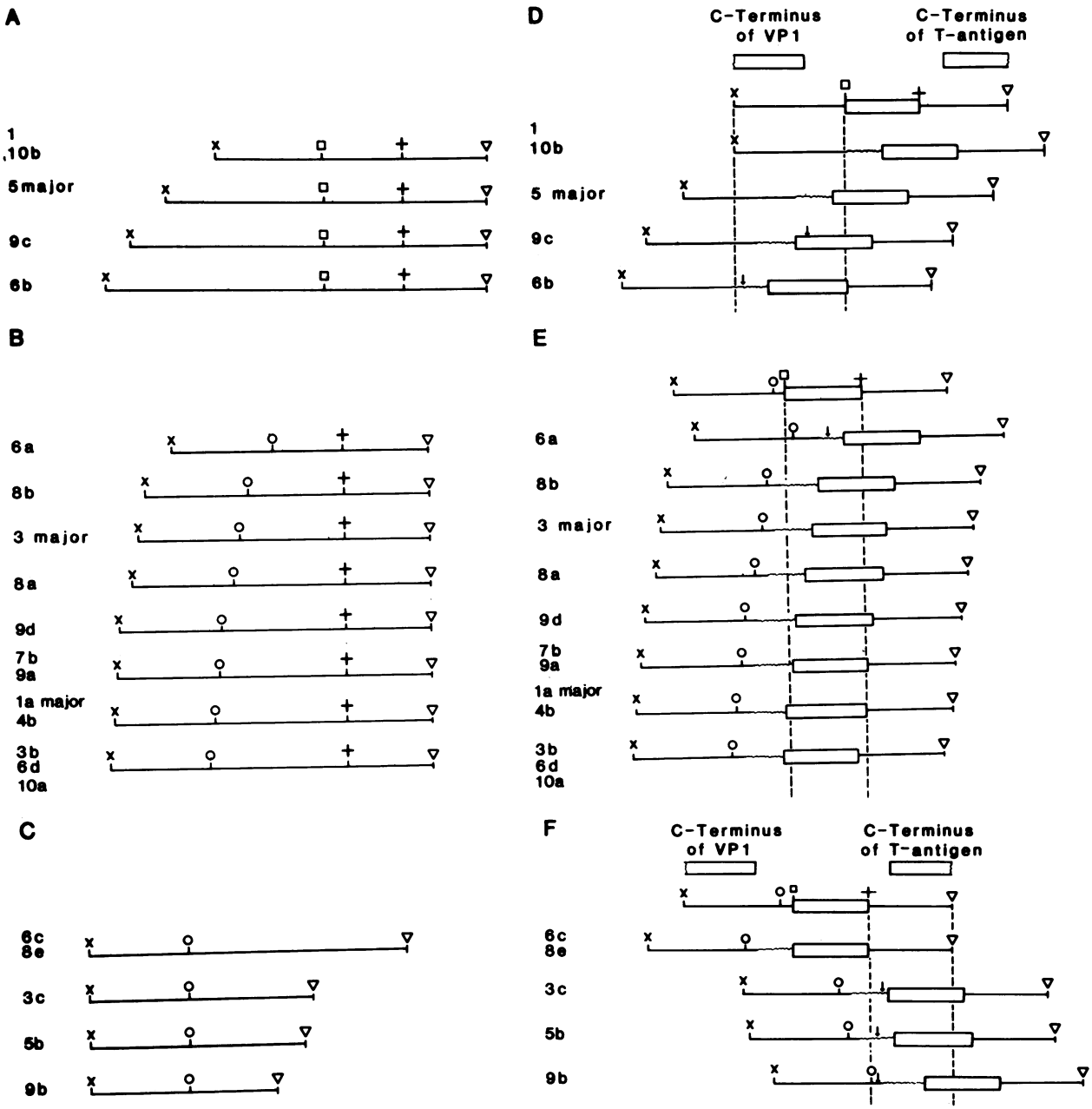


FIG. 8. Physical maps of viral genomes rescued from SV(GT)^2 DNA. (A to C) The relative sizes and restriction maps of class A, GT, and B molecules, respectively. The word "major" in a label indicates that heterogeneity occurred in that sample, and only the molecule present in major proportion is represented here. (D to F) The putative alignments of recombining molecules. At the top of (D to F) is drawn the *Cla-Xba* linked "recipient" parental DNA, and drawn below is a series of molecules representing the probable alignments of each "donor" parental DNA molecule. The dotted vertical lines demarcate the recombination target in the recipient, as determined by analysis with *Cla* and *Xba*. The limits of the sequences encoding the viral protein VP1 and large T antigen are shown in (D) and (F). Vertical arrows mark recombination points in donor molecules, as determined by DNA sequencing. Symbols: (□) poly(GT) DNA; (∿) nonpoly(GT) portion of the GT segment of SV(GT)^2 proximal to the *Bam* site (see Fig. 1); (X) *Bam*HI; (♀) *Alu*I; (+) *Cla*I; (⍑) *Xba*I; and (▽) *Bcl*I.

Bcl sites (Fig. 7B, lanes 3, 7, 11, 14, and 18), and the expected 110-base-pair fragment bounded by *Cla* and *Bcl* sites (Fig. 7C).

Digestion of GT-class molecules with *Mbo* and *Cla* localized the size heterogeneity of these molecules to the region between the *Bam* and *Cla* sites (Fig. 7D). The size heterogeneity was further localized by double digestion with *Cla* plus *Alu* (Fig. 7E). GT-class molecules produced *Cla-Alu* fragments that varied in length in the manner expected from the size heterogeneity in *Bam-Cla* fragments (cf. Fig. 7D and E). The size heterogeneity in GT-class recombinants was not unexpected because poly(GT) segments are internally repetitious and therefore structurally disposed to misalignment.

Physical maps deduced from data in Fig. 5 and 7 are presented in Fig. 8A to C. Figures 8D to F depict the putative alignment of the molecules that recombined to produce the viral genomes diagrammed in Fig. 8A to C. This depiction of recombination alignment is the simplest way to

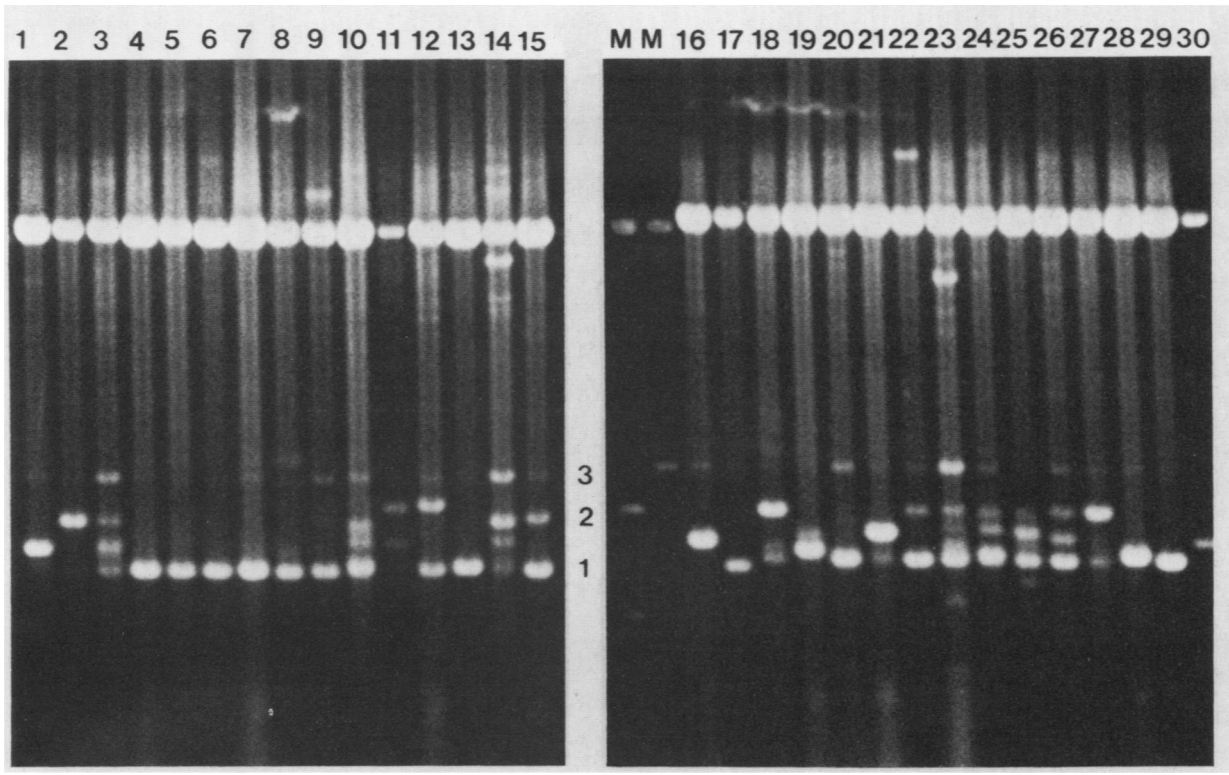


FIG. 9. Viral genomes rescued from SV(AB)^3 DNA. DNA samples were prepared by Hirt extraction of infected cells. DNA was digested with *Pst*I and electrophoresed through 1% agarose; fragments were visualized by ethidium bromide fluorescence. The numbers 1 to 3 in the center space indicate the positions of *Pst*I fragments bearing 1 to 3 copies of the sequence, respectively, between nucleotides 2533 and 2770. Lanes marked (M) contained markers. Lanes 1 through 30 contained DNA from 30 plaques isolated after SV(AB)^3 transfection.

explain the generation of the variety of genomic structures found in rescued viruses. For 11 of the 22 recombinants, sequence analysis (described below) confirmed this interpretation of the restriction data.

Illegitimate deletions were usually not the product of recombination between SV40 and poly(GT) sequences. To determine whether the frequent appearance of nonhomologous recombinants was due to the presence of the poly(GT), an oversized SV40 genome was constructed that contained three tandemly arranged copies of the viral DNA sequence that lies between the *Bam* and *Bcl* cleavage sites. This molecule, SV(AB)^3 (Fig. 3), was the same size as SV(GT)^2 and contained the same tracts of viral sequence homology. SV(AB)^3 was transfected into CVI cells, and viral plaques were rescued as described above. Viral DNAs from 30 plaque isolates were analyzed by digestion with *Pst*I, followed by electrophoresis through agarose. The results of this analysis are shown in Fig. 9 which shows that although most of the plaque isolates rescued from SV(AB)^3 DNA contained *Pst*I fragments of a size consistent with deletion of one or two copies of the *Bam-Bcl* repeat, odd-sized *Pst*I fragments were commonly present (lanes 1, 3, 11, 12, 14, 16, 17, 19, 21, 23–26, 28, and 30). These odd-sized fragments, presumably the product of nonhomologous recombination events, occurred in half the samples, suggesting that even in the absence of poly(GT), oversized viral DNA molecules bearing tandem repeats commonly suffered deletion through illegitimate recombination events. Therefore, the prevalence of illegitimate recombination in the genesis of packageable viral genomes from SV(GT)^2 precursors is for the most part not attributable to the presence of poly(GT).

Consideration of the alignment of molecules depicted in Fig. 8 raised the possibility that the nonhomologous deletions found in the viruses rescued from SV(GT)^2 DNA could have been due to recombination between poly(GT) and SV40 sequences. To test this possibility, five of the six A- and B-class viral genomes that appeared to have been formed by illegitimate recombination were sequenced. All of the viral genomes analyzed were indeed the result of nonhomologous recombination, but only one virus, A9c, was formed by recombination between SV40 and poly(GT) sequences. Four recombination junctions are shown in Fig. 10.

Homologous recombination between two tracts of poly(GT). Six of the 12 GT-class recombinant viruses were sequenced to determine whether the deletions that formed these viral genomes were mediated by homologous recombination between two poly(GT) tracts. Five (GT6d, GT7b, GT8a, GT8b, GT9d) of the six viral genomes analyzed retained perfect tracts of alternating G and T. One virus (GT6a) was clearly formed by nonhomologous recombination. This isolate was the smallest of the 12 GT-class viruses and was the only member of the GT class that was formed by recombination between poly(GT) and non-poly(GT) sequences. The sequence of recombinant GT6a is shown in Fig. 10.

Estimation of the relative proficiencies of poly(GT) and SV40 sequences to recombine homologously. How the raw data shown in Fig. 4 and tabulated in Table 1 relate to the frequency of recombination depends upon whether or not all recombinants are equally able to form a plaque. For those viruses produced by homologous recombination, two lines of evidence argue that differential viability did not influence

FIG. 10. Sequences at illegitimate recombination points. The middle line of each set is the sequence of a recombinant virus. The sequences that participated in the formation of each recombinant are shown above and below each recombinant sequence. Arrows indicate the crossover points. The crossover point for recombinant 5b was ambiguous.

the recovery of viruses produced by the three alternative deletion events shown in Fig. 1. First, viruses of each recombinant class competed well with each other upon mixed infection. CVI cells were infected either singly or mixedly with 100 PFU of a virus representative of each recombinant class. Ten days after infection, viral DNA was prepared by Hirt extraction and analyzed by restriction endonuclease digestion. None of the three viruses outgrew its siblings (data not shown).

A second reason to discount selection as an influence in the recovery of various homologous recombinants is that a mixture of deleted viral genomes produced in *E. coli* yielded virus of each recombinant class in proportion to the amount of DNA of each recombinant class. Upon propagation in *E. coli* with a functional *recA* pathway, pGT^2, the DNA molecule containing the tandem repeats destined for implantation into $SV(GT)^2$ DNA (Fig. 2) suffered frequent deletion of one copy of the repeat. The deletions produced in *E. coli* were of all three types expected to arise by homologous recombination between tandem repeats. The deleted fragments were purified and used to produce virus as described in the legend to Fig. 2. The viral DNA in 30 plaques was analyzed by digestion with *Cla*I and *Xba*I. The distribution of viruses among the three possible classes expected to be produced by homologous recombination was the same as the distribution of deleted DNA fragments in the mixture derived from *E. coli*. This result indicated that the probability of rescue was not influenced by which pair of homologous sequence blocks recombined to produce a packageable viral genome.

Since selection did not favor one homologous recombinant over another, the frequency at which viruses were produced by various homologous recombination events should be a direct indication of the relative propensities of poly(GT) and SV40 sequences to recombine homologously. Although it is clear from restriction endonuclease mapping that only two of the A-class viruses and two of the B-class viruses were formed by homologous recombination, it is more difficult to determine the number of GT-class isolates that were produced by homologous recombination. The size of recombinants is not indicative of mechanism because poly(GT) recombinants would be expected to be variable in size whether or not their formation were mediated by base pairing. Sequence analysis provides some indication of mechanism. Of the six GT-class genomes sequenced, five retained the perfect dinucleotide repeat expected to result from homologous crossing over. The one GT-class genome that was clearly formed by nonhomologous recombination was not informative in regard to the mechanism of GT-by-GT breakage and reunion because only one poly(GT) sequence participated in the recombination event. The fact that five GT-class recombinants appeared to be the result of homologous recombination suggests an upper limit for the contribution of the nonhomologous recombination pathway to the formation of deletions ending in two poly(GT) tracts. Barring the action of a nuclease that might cut poly(GT) tracts only between G and T or only between T and G, there is less than a 10% chance that all of the five sequenced genomes would appear to have been formed by homologous recombination if GT recombination were mediated by the homologous pathway less than 20% of the time. Therefore, it can be concluded with 90% confidence that at least 2 of the 12 GT-class recombinants were formed by homologous recombination and that the frequency of recovery of homologous GT-class viruses was at least as high as the frequency

of recovery of A- or B-class homologous recombinants. Since the target size for each class of homologous recombination event was roughly the same and since all homologous recombinants were equally viable, it follows that, on a per-nucleotide basis, poly(GT) tracts are at least as likely to recombine as two homologous SV40 sequences. An unlikely caveat to this conclusion is that it is possible that poly(GT) recombination appears homologous because a special nuclease cuts these sequences with specificity for cleavage between either GT or TG dinucleotides. However, extensive analysis of illegitimate recombination involving SV40 DNA has not implicated specific cleavage of TG or GT dinucleotides. Rather, trinucleotides of the consensus sequence Py-T-T seem to be involved in nonhomologous recombination of SV40 DNA (4). In bacteria, DNA topoisomerases are thought to mediate nonhomologous recombination (17, 24), and such enzymes are likely candidates for performing such operations in eucaryotes. The dinucleotides GT, CA, and AC are eschewed by mammalian topoisomerase I, and TG is cut rather rarely by this enzyme (6).

An upper limit for the relative homologous recombination proficiency of poly(GT) sequences can also be obtained if 11 GT-class viruses are assumed to have been homologous recombinants. One would expect to isolate this proportion of GT-class recombinants if the 85 base pairs of poly(GT) had presented a target for homologous recombination that was 2.75 times larger than the target presented by the 237 base pairs of SV40 DNA available to mediate deletion via homologous recombination. If the 85 base pairs of poly(GT) were to have behaved as if they were a recombination target 651 base pairs long, it would follow that, on a per-nucleotide basis, recombination between poly(GT) tracts was 7.7 times more likely (651 divided by 85) than recombination between homologous sequence tracts composed of SV40 DNA.

DISCUSSION

The data presented in this report indicate that recombination between poly(GT) repeats readily occurs in SV40-infected cells. The finding that viruses rescued from $SV(GT)^2$ DNA were commonly the product of recombination between poly(GT) segments indicates that the previously described stability of poly(GT) in a viable variant of SV40 that carries a poly(GT) insertion (36) was not due to a strict prohibition of recombination between poly(GT) sequences. The apparent stability of poly(GT) in SVGT DNA may reflect the infrequency of deletion and the relatively small gain in replicative proficiency associated with deletion of most, but not all, of the poly(GT) insertion.

The size heterogeneity of poly(GT) tracts left in viral genomes formed via GT-by-GT recombination indicates that recombination between two tracts of poly(GT) is usually unequal. Two lines of evidence argue that the size heterogeneity of poly(GT) traits is due to unequal recombination and not due to secondary deletions of poly(GT) after the recombination events that deleted one copy of the reiterated *Bam-Bcl* fragment. First, heterogeneity of poly(GT) occurred only in molecules formed by recombination between two poly(GT) tracts (Fig. 7). None of the A-class recombinants lost nucleotides from the poly(GT) segment flanked by *Xba* and *Cla* sites. Second, poly(GT) has been shown to be stably maintained in SV40 genomes of packageable size (36).

The poly(GT) tracts resultant from unequal poly(GT) recombination were distributed by size over a continuum ranging between 40 and 80 base pairs. The one GT-class recombinant virus that deviated from this continuous size distribution, isolate GT6a, was found to be the product of recombination between a poly(GT) tract and nonpoly(GT) DNA. This size range is what would be expected if homologous recombination of poly(GT) sequences were to require at least 40 base pairs of sequence homology. Forty base pairs has been previously determined to be the mimimum amount of homology required to mediate efficient homologous recombination in T4-infected *E. coli* (32). The variation in size of poly(GT) tracts left in GT-class recombinants could also reflect the action of the nonhomologous deletion pathway.

$SV(GT)^2$ DNA commonly produced viruses that were the product of nonhomologous recombination. Although SV40 DNA is known to recombine rather promiscuously both with itself and with totally heterologous DNA molecules (37, 40, 42), it was still surprising to find that progeny molecules arising by deletions involving either of the two SV40 sequence tracts were usually the product of nonhomologous recombination. Comparison of $SV(GT)^2$ and a similar DNA molecule that lacked polyGT, $SV(AB)^3$, showed that plaques produced by $SV(AB)^3$ were as likely to contain viruses formed by nonhomologous recombination as plaques produced by $SV(GT)^2$ DNA. This suggests that the incidence of nonhomologous recombination in $SV(GT)^2$ DNA could in large part be ascribed to the general promiscuity of SV40 recombination. Furthermore, although the alignment of recombining molecules probably placed poly(GT) sequences in a position to undergo nonhomologous recombination with non-poly(GT) sequences, sequence analysis showed that nonhomologous deletions were usually not the result of the direct involvement of poly(GT) in illegitimate breakage and reunion events.

The prevalance of nonhomologous recombination precluded the exact determination of the relative proficiency of poly(GT) sequences to undergo homologous recombination, because an indeterminate fraction of the viral genomes generated by nonhomologous deletions in $SV(GT)^2$ DNA would escape detection in the virus rescue protocol and because it is difficult to assess the proportional contribution of the nonhomologous recombination pathway to poly(GT) recombination. However, the data allow the estimation of the maximum enhancement of homologous recombination frequency that may be associated with poly(GT) DNA. Depending upon the fraction of GT-class viruses that were formed via homologous recombination between two poly(GT) tracts, an 85-base-pair poly(GT) sequence was as much as eight times more likely to recombine homologously than 85 base pairs of SV40 DNA. Such a relatively high recombination propensity could presumably be a manifestation of the unusual structure of poly(GT) DNA. One possibility is that, as they do in vitro, poly(GT) sequences assume a left-handed helical conformation in the cell. In consideration of this possibility, it is interesting that the noninterwound recombination intermediates fostered by the in vitro action of the *Ustilago* RecI protein have been shown to bind antibodies specific for DNA in the Z conformation (20). Also, poly(GC) DNA in the Z form has been shown to bind RecI protein in vitro (19). Alternatively, recombination between poly(GT) sequences may have nothing to do with Z DNA but simply reflect the fact that the sequence repeats itself every 2 base pairs. Any dinucleotide repeat sequence may be more likely to form a homologously matched four-strand recombination intermediate. The in vitro action of *E. coli* RecA protein entails a homology-searching step (8). Such searches would be expected to succeed more quickly, and therefore perhaps more often, when they involve two

sequences composed of a dinucleotide repeat. The pertinence of the functional properties of the *E. coli* RecA protein to the mechanism of homologous recombination in eucaryotes is strongly supported by the finding of similar in vitro activities expressed by the RecI protein of the fungus *Ustilago maydis* (20).

The behavior of poly(GT) sequences in $SV(GT)^2$ suggests that a chromosomal poly(GT) sequence of a given length could recombine homologously up to eight times more frequently than a more complex sequence of the same length. In this sense, poly(GT) sequences could be regarded as hot spots for recombination. However, because cellular poly(GT) sequences are in general quite short, 10 to 50 base pairs, it seems unlikely that they actually function as hot spots for reciprocal recombination between homologous or closely related genes. Even a 10-fold enhancement of recombination would only increase the recombination of a 50-base-pair poly(GT) sequence to a level equal to that expected to occur between 500 base pairs of more complex sequence. Therefore, although it seems that the magnitude of recombination enhancement of poly(GT) sequences might be sufficient to offset their small size to the point of rendering poly(GT) sequences as likely to recombine as surrounding tracts of more complex sequence, small cellular poly(GT) sequences probably should not be regarded as potential hot spots for reciprocal recombination between two homologous or very closely related genes.

ACKNOWLEDGMENTS

I thank Jim Newman for technical assistance and Keith Murphy and Holly Schmidt for help in preparing the manuscript.

This work was supported by grant MV175 from the American Cancer Society and by Public Health Service grant GM-31789 from the National Institutes of Health.

LITERATURE CITED

1. **Brockman, W. W., M. W. Gutai, and D. Nathans.** 1975. Evolutionary variants of simian virus 40: characterization of cloned complementary virus. Virology **66**:36–52.
2. **Brockman, W. W., T. N. H. Lee, and D. Nathans.** 1973. The evolution of new species of viral DNA during serial passage of SV40 at high multiplicity. Virology **54**:384–397.
3. **Buchman, A. R., L. Burnett, and P. Berg.** 1981. The SV40 nucleotide sequence, p. 799–841. *In* J. Tooze (ed.), The molecular biology of tumor viruses, part 2, 2nd ed., revised. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
4. **Bullock, P., W. Forrester, and M. Botchan.** 1984. DNA sequence studies of simian virus 40 chromosomal excision and integration in rat cells. J. Mol. Biol. **174**:55–84.
5. **Caulcott, E.** 1973. Significance tests. Routledge and Kegan Paul Ltd., London.
6. **Edwards, K. A., B. D. Halligan, J. L. Davis, N. L. Nivera, and L. F. Liu.** 1982. Recognition sites of eukaryotic DNA topoisomerase I: DNA nucleotide sequencing analysis of topo I cleavage sites on SV40 DNA. Nucleic Acids Res. **10**:2565–2576.
7. **Flanagan, J. C., M. P. Lefranc, and T. H. Rabbitts.** 1984. Mechanisms of divergence and convergence of the human immunoglobulin α1 and α2 constant region gene sequences. Cell **36**:681–688.
8. **Gonda, D. K., and C. M. Radding.** 1983. By searching processively RecA protein pairs DNA molecules that share a limited stretch of homology. Cell **34**:647–654.
9. **Hamada, H., and T. Kakunaga.** 1982. Potential Z-DNA forming sequences are highly dispersed in the human genome. Nature (London) **298**:396–398.
10. **Hamada, H., M. G. Petrino, and T. Kakunaga.** 1982. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. U.S.A. **79**:6465–6469.
11. **Hamada, H., M. G. Petrino, T. Kakunaga, M. Seidman, and B. D. Stollar.** 1984. Characterization of genomic poly(dT-dG) · poly(dC-dA) sequences: structure, organization, and conformation. Mol. Cell. Biol. **4**:2610–2621.
12. **Hamada, H., M. Seidman, B. H. Howard, and C. M. Gorman.** 1984. Enhanced gene expression by the poly(dT-dG) · poly(dC-dA) sequence. Mol. Cell. Biol. **4**:2622–2630.
13. **Hanahan, D.** 1983. Studies of transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
14. **Haniford, D. B., and D. E. Pullyblank.** 1983. Facile transition of poly [d(TG)-d(CA)] into a left-handed helix in physiological conditions. Nature (London) **302**:632–534.
15. **Hirt, B.** 1967. Selective extraction of polyoma virus DNA. J. Mol. Biol. **26**:365–369.
16. **Hochtl, J., and H. G. Zachan.** 1983. A novel type of aberrant recombination in immunoglobulin genes and its implications for V-J joining mechanism. Nature (London) **302**:260–262.
17. **Ikeda, H., K. Aoki, and A. Naito.** 1982. Illegitimate recombination mediated *in vitro* by DNA gyrase of *Escherichia coli*: structure of recombinant DNA molecules. Proc. Natl. Acad. Sci. U.S.A. **79**:3724–3728.
18. **Kim, S., M. Davis, E. Sinn, P. Patten, and L. Hood.** 1981. Antibody diversity: somatic hypermutation of rearranged V_H genes. Cell **27**:573–581.
19. **Kmiec, E. B., K. J. Angelides, and W. L. Holloman.** 1985. Left handed DNA and the synaptic pairing reaction promoted by Ustilago RecI protein. Cell **40**:139–145.
20. **Kmiec, E. B., and W. K. Holloman.** 1984. Synapsis promoted by Ustilago RecI protein. Cell **36**:593–598.
21. **Lai, C. J.** 1980. Mapping SV40 mutants by marker rescue. Methods Enzymol. **65**:811–816.
22. **Liu, C.-P., P. W. Tucker, J. F. Mushinski, and F. R. Blattner.** 1980. Mapping of heavy chain genes for mouse immunoglobulins M and D. Science **209**:1348–1353.
23. **Maniatis, T., E. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
24. **Marvo, S. L., S. R. King, and S. R. Jaskunas.** 1983. Role of short regions of homology in intermolecular illegitimate recombination events. Proc. Natl. Acad. Sci. U.S.A. **80**:2452–2456.
25. **Maxam, A., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
26. **Miesfeld, A., M. Krystal, and N. Arnheim.** 1981. A member of a new repeated sequence family which is conserved throughout eucaryotic evolution found between the human and β globin genes. Nucleic Acids Res. **9**:5931–5947.
27. **Nishioka, Y., and P. Leder.** 1979. The complete sequence of a chromosomal mouse α-globin gene reveals elements conserved throughout vertebrate evolution. Cell **18**:875–882.
28. **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.
29. **Qasba, P. K., and S. K. Safaya.** 1984. Similarity of the nucleotide sequences of rat α-lactalbumin and chicken lysozyme genes. Nature (London) **308**:377–380.
30. **Rubnitz, J., and S. Subramani.** 1984. The minimum amount of homology required for homologous recombination in mammalian cells. Mol. Cell. Biol. **4**:2253–2256.
31. **Shen, S.-H., and O. Smithies.** 1982. Human globin $\psi\beta_2$ is not a globin-related sequence. Nucleic Acids Res. **10**:7809–7818.
32. **Singer, B. D., L. Gold, P. Gauss, and D. H. Doherty.** 1982. Determination of the amount of homology required for recombination in bacteriophage T_4. Cell **31**:25–33.
33. **Slightom, J. L., A. E. Blechl, and O. Smithies.** 1980. Human fetal G_γ and A_γ-globin genes: complete nucleotide sequences suggest that DNA can be exchanged between these duplicated genes. Cell **21**:627–638.
34. **Stanton, L. W., R. Watt, and K. Marcu.** 1983. Translocation, breakage, and truncated transcripts of c-myc oncogene in murine plasmacytomas. Nature (London) **303**:401–406.
35. **Stringer, J. R.** 1982. DNA sequence homology and chromo-

somal deletion at a site of SV40 DNA integration. Nature **296**:363–366.
36. **Stringer, J. R.** 1985. Construction of a viable simian virus 40 variant that carries a poly[d(GT) · d(CA)] insertion. J. Virol. **53**:698–701.
37. **Subramani, S., and P. Berg.** 1983. Homologous and nonhomologous recombination in monkey cells. Mol. Cell. Biol. **3**:1040–1052.
38. **Swanstrom, R., R. C. Parker, H. E. Varmus, and J. M. Bishop.** 1983. Transduction of a cellular oncogene: the genesis of Rous sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **80**:2519–2523.
39. **Vasicek, T. J., B. A. McDevitt, M. W. Freeman, B. J. Fennick, G. N. Hendy, J. T. Potts, Jr., A. Rich, and H. M. Kronenberg.** 1983. Nucleotide sequence of the human parathyroid hormone gene. Proc. Natl. Acad. Sci. U.S.A. **80**:2127–2131.
40. **Wake, C. T., T. Gudewicz, T. Porter, A. White, and J. H. Wilson.** 1984. How damaged is the biologically active subpopulation of transfected DNA? Mol. Cell. Biol. **4**:387–398.
41. **Wake, C. T., and J. H. Wilson.** 1980. Defined oligomeric SV40 DNA: a sensitive probe of general recombination in somatic cells. Cell **21**:141–148.
42. **Winocour, E., and I. Keshet.** 1980. Indiscriminant recombination in simian virus 40-infected monkey cells. Proc. Natl. Acad. Sci. U.S.A. **77**:4861–4865.