JOURNAL OF VIROLOGY, May 1986, p. 393–401
0022-538X/86/050393-09$02.00/0

Vol. 58, No. 2

Formation of Deletions after Initiation of Simian Virus 40 Replication: Influence of Packaging Limit of the Capsid

XIU-BAO CHANG AND JOHN H. WILSON*

Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, Texas 77030

Received 3 September 1985/Accepted 4 February 1986

Transfected DNA is frequently broken and rejoined in mammalian cells by recombination processes that depend on minimal nucleotide sequence homology. Although measurements of breakage and joining account reasonably well for the frequent formation of deletions during transfection, they are inadequate to explain the high frequency of deletion formation by simian virus 40 (SV40) genomes that are slightly larger than the packaging limit of the capsid. To investigate this anomaly, we constructed and transfected into CV-1 cells a series of modified SV40 genomes containing 136, 284, 460, and 656 extra base pairs in the intron of the gene encoding T antigen. These experiments indicate that the effective packaging limit of an SV40 capsid lies between 284 and 460 extra base pairs. Further analysis of these transfections suggests that molecules just above the effective packaging limit may be encapsidated and transmitted between cells at low efficiency, thereby allowing multiple rounds of replication and multiple opportunities to generate and package genomes that contain deletions. The junctional sequences in several such deletions were determined; they were similar to the junctions in deletions that were formed before replication began, suggesting that the enzymatic machinery responsible for both types of deletion may be similar.

Nonhomologous recombination can be studied conveniently in the intron of the simian virus 40 (SV40) T-antigen gene because the intron is dispensable for lytic growth. Since no particular sequence must be reconstructed to allow expression of T antigen and subsequent plaque formation, recombinants in the intron are viable, making a spectrum of products readily available for analysis. The distribution of types of recombinant within this spectrum reflects the enzymatic activities of the cell and can be used to deduce certain characteristics of the overall recombination process.

Using SV40 genomes containing extra DNA in the intron, we have demonstrated that the primary mechanism of nonhomologous recombination in transfected DNA involves breakage of the input DNA followed by joining of ends (17). Although neither the breakage nor the joining step is well understood, we have shown that input DNA in the nucleus is broken about once per 5 to 15 kilobases and that sticky, blunt, and mismatched ends are joined with nearly equal efficiency (14). In addition, the types and frequencies of the various recombinants produced by genomes with mismatched ends are consistent with the measured parameters for breakage and joining (7).

Although the measured frequencies of breakage and joining account reasonably well for the infectivity of most oversized genomes of SV40, they do not predict the high efficiency of plaque formation that is observed for genomes just above the packaging limit for the viral capsid. For example, SV40 genomes with 452 extra base pairs in the intron must be above the packaging limit as judged by the absence of full-length genomes in plaques that arise after DNA transfection (14). If the deleted genomes found in these plaques arose by breakage and rejoining in transfected DNA, plaques should be produced at about 5% of wild-type efficiency (7, 17). However, genomes with 452 extra base pairs yielded plaques at 68% of the efficiency of wild-type genomes, suggesting that a process other than breakage and joining in transfected DNA might be responsible.

* Corresponding author.

The anomaly at the packaging limit raises the possibility that the packaging process itself might be directly or indirectly involved. To understand the basis for the anomaly, we constructed a series of different-length SV40 genomes that span the packaging limit, ranging in size from 136 extra base pairs, which is packageable, to 656 extra base pairs, which is not (14). Linear and circular versions of these genomes were transfected into CV-1 monkey kidney cells. Their relative infectivities were measured, the time course for appearance of deleted genomes was determined, and genomes from individual plaques were characterized. These experiments show that the effective packaging limit for an SV40 capsid lies between 284 and 460 extra base pairs. In addition, they indicate that the underlying deletion process occurs after DNA replication has been initiated and suggest that a rare deletion process is being amplified by the packaging properties of the viral capsid.

MATERIALS AND METHODS

Cells and viruses. Procedures for growth of the established monkey kidney CV-1 cell line have been described elsewhere (18). Viral titers were determined by plaque assay on CV-1 cells as previously described (18). Wild-type SV40 is strain Rh911.

DNA preparation. SV40 DNA and plasmid DNA were prepared as described previously (2, 17). SV40 and plasmid DNA preparations used for constructions and transfections were routinely labeled in vivo with [^{3}H]thymidine. Specific activities were determined and used to adjust DNA concentrations as appropriate for each experiment.

Construction of oversized SV40 genomes. Restriction enzymes, T4 DNA ligase, T4 DNA polymerase, DNA polymerase I (Klenow fragment), and T4 polynucleotide kinase were purchased from New England BioLabs, Inc., or Boehringer Mannheim Biochemicals and used according to the recommendations of the suppliers. The construction scheme for creating SV40 genomes with extra base pairs is shown in Fig. 1. After digestion of the parent plasmid, pJJ4, with the indicated restriction enzymes, the ends were blunted to

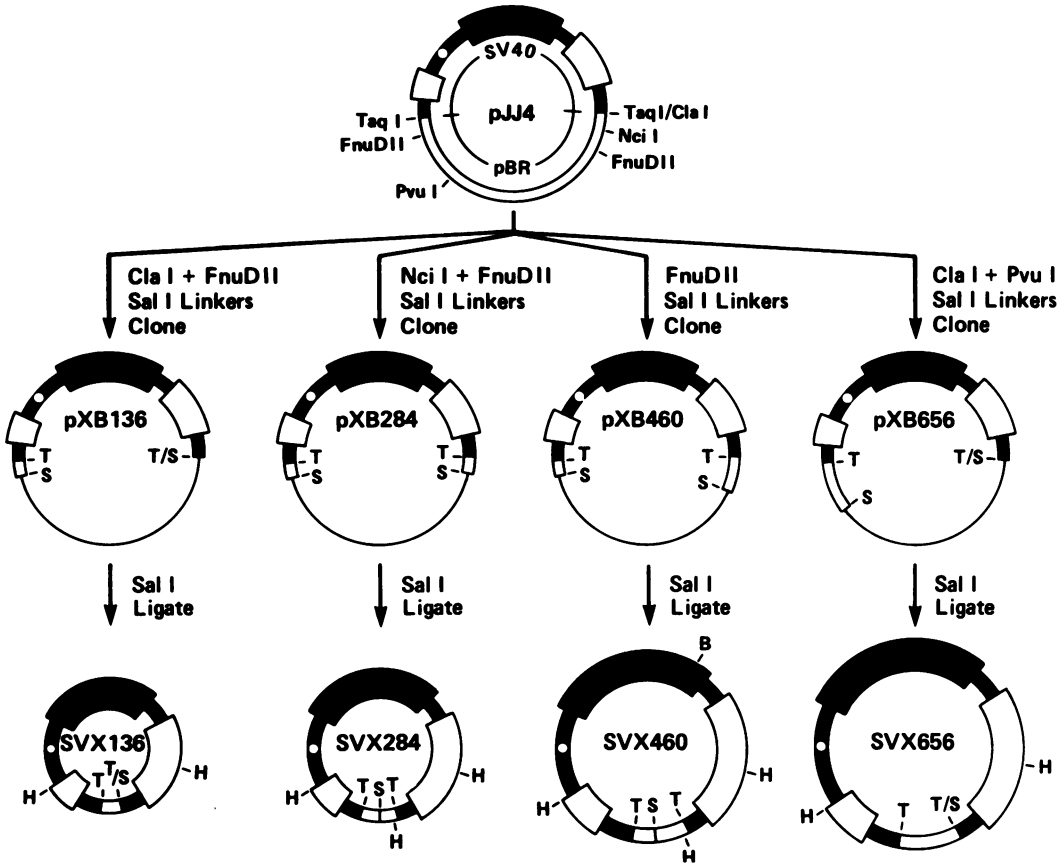


FIG. 1. Construction of SVX genomes. The parent plasmid, pJJ4, contains SV40 inserted through its unique *Taq*I site into the *Cla*I site of pBR322. To generate the pXB series of plasmids, pJJ4 was restricted with the indicated enzymes, the ends were blunted, *Sal*I linkers were added, and the modified genome was cloned into the *Sal*I site of pBR322. SVX genomes were derived from the pXB plasmids by digestion with *Sal*I. Circular SVX genomes were produced by ligation of *Sal*I-cleaved linear genomes at a low concentration. Restriction sites used in construction and analysis are as follows: T (*Taq*I), S (*Sal*I), B (*Bam*HI), and H (*Hin*dIII). Only a subset of *Hin*dIII sites is shown.

accept *Sal*I linkers. Ends with 5′ protrusions were filled in with DNA polymerase I (Klenow fragment), and ends with 3′ protrusions were blunted by incubation with T4 DNA polymerase. After the addition of *Sal*I linkers, the linear molecules were purified from agarose gels after dissolution in saturated KI (9, 19), digested with *Sal*I, and cloned into the *Sal*I site of pBR322.

DNA transfection. DNA transfections were essentially by the method of McCutchen and Pagano (4) with DEAE-dextran (molecular weight, 500,000; Pharmacia Fine Chemicals, Inc.) at 500 μg/ml in a volume of 0.3 ml. All DNA plaque assays were carried out on 60-mm plastic petri plates that contained freshly confluent or slightly subconfluent cell monolayers as described previously (16). Under the conditions of transfection used in these experiments, more than 95% of all plaques should be initiated by single molecules (7, 15; U. Weiss and J. H. Wilson, unpublished data). After transfection, the monolayers were overlaid with agar and incubated at 37°C as in a standard plaque assay. Well-isolated plaques were picked for analysis.

Miniwell preparation of DNA. Confluent CV-1 cells in 96-well microtiter plates were infected with picked plaque suspensions. Viral DNA was labeled in vivo by the addition of $^{32}P_i$ in phosphate-free medium and harvested as described previously (17).

Southern blot analysis. Viral DNA from Hirt supernatants (2) was treated with RNase, phenol extracted twice, butanol extracted once, and ethanol precipitated once with sodium acetate. The DNA was then suspended in 10 mM Tris–1 mM EDTA (pH 7.4) and ethanol precipitated twice with ammonium acetate. After digestion, restriction fragments were separated by electrophoresis on 1% agarose gels. Gels were soaked twice in 0.5 N NaOH for 30 min at room temperature, rinsed three times with deionized water, and neutralized by being soaked twice in 0.5 M Tris hydrochloride (pH 7.0) for 30 min at room temperature. After equilibration of the gel in TEA buffer (10 mM Tris, 5 mM sodium acetate, 0.5 mM EDTA [pH 7.8]) for 30 min at room temperature, the DNA was transferred to Zetabind membranes (AMF Cuno) by overnight blotting with TEA buffer (10). The membranes were then baked for 2 h at 80°C in a vacuum oven and washed for 1 h at 65°C in 0.5× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) containing 0.5% sodium dodecyl sulfate (SDS). Prehybridizations and hybridizations were done overnight at 65°C in 50 mM phosphate buffer (pH 7.0) containing 5× SSC; 0.2% each bovine serum albumin, Ficoll, and polyvinylpyrrolidone; 1% SDS; and 500 μg of sonicated salmon sperm DNA per ml. Nick-translated pXB284 at 10^6 counts/ml was used as the hybridization probe (6). After hybridization, membranes were washed twice at room temperature in 3× SSC containing 1% SDS, followed by two washes at 65°C in 0.5× SSC–1% SDS–0.1% PP_i. The membranes were then dried, and the bands were visualized by autoradiography.

Nucleotide sequencing. Nucleotide sequences were determined by the dideoxy method of Sanger (8) as modified for double-stranded DNA (20). Purified viral DNA was hybridized to an excess of a synthetic primer and then incubated with Moloney mouse leukemia virus reverse transcriptase (Bethesda Research Laboratories, Inc.) at 42°C for 20 min in a reaction mixture containing all four deoxynucleoside triphosphates, ^{35}S-dATP (Amersham Corp.), and one of four dideoxynucleoside triphosphates. Samples were then heat denatured and subjected to electrophoresis on a denaturing polyacrylamide gel. Bands were visualized by autoradiography.

RESULTS

Construction of SV40 genomes containing extra base pairs. To construct a series of oversized SV40 genomes, we cleaved the plasmid pJJ4, which contains pBR322 inserted in the intron of the T-antigen gene, with combinations of restriction enzymes that do not cut SV40 DNA. We then added *Sal*I linkers to the fragments containing the SV40 genome and recloned these fragments into the *Sal*I site in pBR322 (Fig. 1). Cleavage of the resulting pXB plasmids with *Taq*I liberates a wild-type SV40 genome, whereas cleavage with *Sal*I liberates an oversized SV40 genome containing extra base pairs in the intron. The circular forms of the oversized SVX genomes are shown at the bottom of Fig. 1. The number following the SVX designation indicates the number of extra base pairs relative to a wild-type genome.

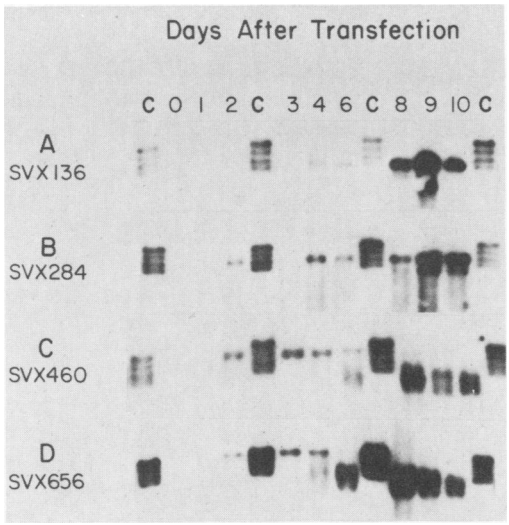


FIG. 2. Time course of infection with oversized linear genomes. (A) SVX136; (B) SVX284; (C) SVX460; (D) SVX656. Oversized genomes, which had been linearized by digestion with *Sal*I, were transfected into CV-1 cells at 10 ng per 60-mm plate. Immediately after transfection (day 0) and at various later times, viral DNA was harvested from two plates. The viral DNA was then digested with *Bam*HI, fractionated by electrophoresis on 0.7% agarose gels, transferred to Zetabind membranes, hybridized with nick-translated pXB284, and visualized by autoradiography. The two fragments generated by cleaving parent linear DNA ran off the bottom of these gels; the indicated bands had migrated nearly 20 cm. Numbers at the top indicate days after transfection. Sample volumes were adjusted to give more equivalent hybridization signals (days 0 and 1, 30 μl; days 2 to 4, 10 μl; day 6, 3 μl; day 8, 1 μl; day 9, 0.5 μl; day 10, 0.25 l). The control (C) lanes contained a mixture of linear SVX136, SVX284, SVX460, and SVX656 genomes. Loss of DNA in the samples from days 2 and 3 for SVX136 and from day 3 for SVX284 presumably occurred during preparation.

TABLE 1. Infectivity of oversized SVX genomes

Genome	Form[a]	Relative infectivity (%)[b]
SVX136	C	76 ± 19
	L_S	68 ± 30
SVX284	C	77 ± 30
	L_S	44 ± 26
SVX460	C	49 ± 14
	L_S	32 ± 10
	L_B	72 ± 28
SVX656	C	11 ± 4
	L_S	12 ± 10

[a] Circular (C) and linear (L) forms of the SVX genomes were purified from agarose gels and transfected into CV-1 cells. Cleavage at the *Sal*I site (L_S) splits the T-antigen gene and positions the added nucleotides at the termini; cleavage at the *Bam*HI site (L_B) splits the late coding block and positions the extra nucleotides internally (Fig. 1).

[b] Data points are the average of five experiments. In each experiment, nicked circular and linear SVX genomes and the corresponding wild-type forms, which were derived from the same parent plasmid after digestion with *Taq*I, were plaque assayed at 1 ng per plate on five plates. The infectivity of each SVX genome was normalized to the infectivity of the corresponding wild-type genome and is expressed in the table as relative infectivity.

Packaging limit of SV40. To assess the packaging limit of an SV40 capsid, we measured the retention of the transfected, parent-length genome through multiple cycles of infection. Linear forms of SVX136, SVX284, SVX460, and SVX656 were prepared by digestion with *Sal*I and then purified from agarose gels. These genomes were transfected into CV-1 cells, in which the SV40 growth cycle is 2 to 3 days (13), and viral DNA was harvested over a 10-day period. The viral DNA was restricted with *Bam*HI, which linearizes circular forms but splits the parental linear DNA into two fragments, and displayed by blot hybridization after electrophoresis on an agarose gel. At late times during infection, parent-length genomes were the majority species in SVX136 and SVX284 transfections but were undetectable in SVX460 and SVX656 transfections (Fig. 2). These results show that the packaging limit for an SV40 capsid lies between 284 and 460 extra base pairs.

Infectivity of oversized genomes. Because the two smaller genomes (SVX136 and SVX284) can be packaged, whereas the two larger genomes (SVX460 and SVX656) cannot, one might expect the infectivities of the small genomes to be high and roughly equal and the infectivities of the large genomes to be low and roughly equal. Linear and circular forms of SVX136, SVX284, SVX460, and SVX656 were purified from agarose gels and then transfected into CV-1 monkey cells; their infectivities were assayed by plaque formation. In each case, the infectivity of the SVX genome was normalized to the infectivity of the corresponding wild-type form liberated from the same pXB parent plasmid by digestion with *Taq*I. The results of these transfections are listed in Table 1 and displayed in Fig. 3. The data for circular genomes and *Sal*I-cleaved linear genomes (Fig. 3, solid lines) parallel our previous infectivity measurements for oversized linear genomes, which were assayed in a mixture with pBR322 DNA (Fig. 3, dashed line). Within the accuracy of the assay, linear genomes (open symbols) and circular genomes (filled symbols) had the same infectivity.

The infectivities of SVX136 and SVX284 were similar and high, as expected, and the infectivity of SVX656 was low, although 2 to 3 times higher than predicted from our measured parameters for breakage and rejoining. However, the infectivity of SVX460 was much higher, at least 10-fold higher, than predicted. This efficient plaque formation by

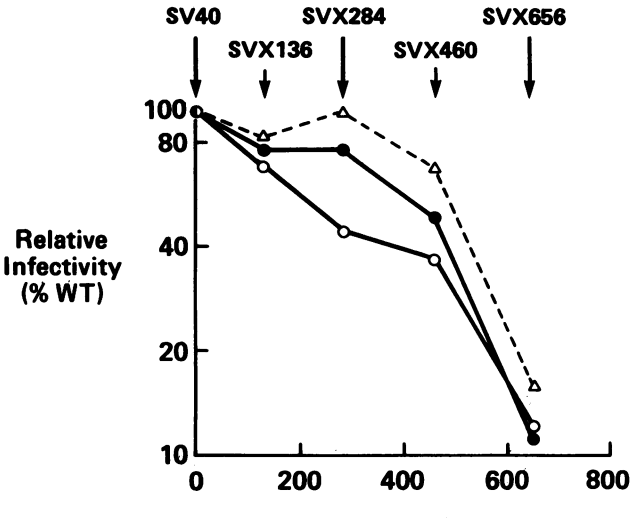


FIG. 3. Infectivity of oversized SVX genomes. Relative infectivities for circular (filled circles) and *Sal*I-cleaved linear (open circles) SVX genomes are from Table 1. Open triangles represent previous data (17) derived from genomes liberated from pJJ4 by the same restriction enzymes used in the construction of the SVX genomes.

genomes just above the packaging limit is the anomaly under investigation. In some way deletions are generated with extremely high efficiency during infection by SVX460; they are also generated at a lower efficiency during infection by SVX656.

Deletions are formed after end joining. If the anomaly at the packaging limit were due to breakage and joining events in the transfected DNA molecules, we would expect linear molecules just above the packaging limit to be significantly more infectious than circular molecules. Studies with shuttle vectors have demonstrated that linear molecules with ends near the target sequence generate deletions during transfection 10 times more frequently than do circles (5). (Presumably, a linear genome with an end near the target sequence need suffer only one break to generate a deletion, whereas a circular molecule must suffer two breaks.) The infectivities of linear and circular forms were about the same (Fig. 3). These results suggest that the high efficiency of deletion formation in transfections with SVX460 was not due to breakage and rejoining events in the transfected DNA.

To address this question more directly, we isolated several plaques from each transfection and characterized the genomes they contained. As illustrated schematically in Fig. 4, the status of the *Sal*I site in the isolated genomes allows one to decide whether deletions occurred before or after end joining. If deletions occurred before end joining, *Sal*I$^+$ plaques should not be generated at an appreciable frequency from *Sal*I-linearized genomes, since removal of DNA from either terminus would destroy the *Sal*I site. By contrast, if deletion occurred after end joining, some *Sal*I$^+$ plaques would be expected, because not all subsequent deletions would remove the *Sal*I site.

Virus from picked plaques were amplified in miniwells in the presence of $^{32}PO_4$, viral DNA was purified and restricted with *Hin*dIII and *Sal*I, and the resulting fragments were separated by electrophoresis on polyacrylamide gels and visualized by autoradiography. Representative examples of this analysis are shown in Fig. 5. If either one of the diagnostic *Sal*I-*Hin*dIII fragments was present (arrows in Fig. 5), the plaque was classified as *Sal*I$^+$. All transfections yielded a significant number of *Sal*I$^+$ plaques (Table 2 and Fig. 6), indicating that most of the SVX460 deletions and many of the SVX656 deletions were generated after end joining.

Compared with the corresponding circular genomes, all linear SVX genomes generated a lower frequency of *Sal*I$^+$

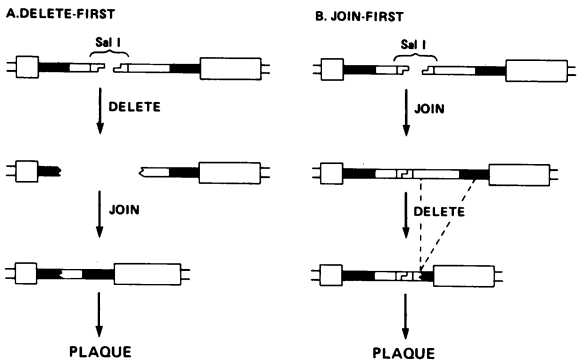


FIG. 4. Two pathways for deletion of DNA from oversized SVX genomes. The intron region of *Sal*I-linearized SVX460 is illustrated at the top of each pathway. The large open boxes represent the exons of the gene for T antigen. The open segments within the intron represent the extra nucleotides derived from pBR322; the filled segments within the intron represent SV40 sequences. The consequences of deletion before (A) or after (B) end joining are illustrated (see text).

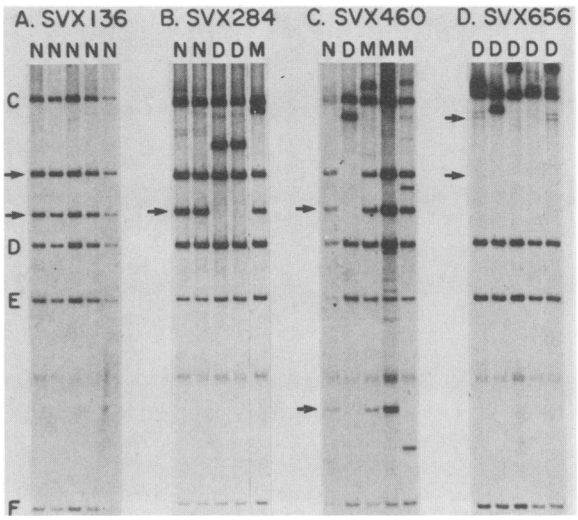


FIG. 5. Plaques generated during transfection of oversized SVX genomes. Plaques in panels A, C, and D arose after transfection with circular forms; plaques in panel B arose after transfection with *Sal*I-cut linears. ^{32}P-labeled viral DNA was prepared from individual plaques by the miniwell method (see Materials and Methods). The restriction fragments produced by digestion with *Hin*dIII and *Sal*I were separated by electrophoresis on 5% polyacrylamide gels and visualized by autoradiography. The positions of the C, D, E, and F fragments generated by *Hin*dIII cleavage of wild-type SV40 are indicated. Plaques were classified as nondeletion (N), deletion (D), or mixed (M) according to criteria described in the text. The two restriction fragments diagnostic for the *Sal*I site in each SVX genome are indicated with arrows; if either fragment was present, the plaque was classified as *Sal*I$^+$. The second fragment in panel B runs well below fragment F; it was missing in the two lanes marked D.

TABLE 2. Plaque types generated by oversized genomes

Genome	Form[a]	Total no. of plaques	% T for[b]:			
			SalI$^+$	Non-deletion	Deletion	Mixed
SVX136	C	47	98	98	2	0
	L_S	36	67	72	25	3
SVX284	C	53	95	74	6	21
	L_S	24	69	67	17	17
SVX460	C	96	91	23	13	65
	L_S	56	62	18	16	66
	L_B	32	81	9	16	75
SVX656	C	17	24	0	82	18
	L_S	22	17	0	91	9

[a] Linear and circular forms are as described in Table 1, footnote *a*. The distributions of plaques derived from transfections with supercoiled and nicked circular genomes were indistinguishable and have been added together in the table.

[b] Plaques were classified as *SalI*$^+$ or *SalI*$^-$ and also as nondeletion, deletion, or mixed according to criteria described in the text. The number of plaques in each category is expressed as a percentage of total plaques (%T).

plaques. These differences presumably arose from minor modifications of the termini. In previous studies with linear genomes with sticky or blunt ends, we demonstrated that 12 and 20%, respectively, of the ends were modified by removal or addition of fewer than 25 base pairs (7, 14). Analysis of genomes present in the *SalI*$^-$ plaques derived from linear SVX284 genomes revealed the same sort of minor end modification (two examples, marked D, are shown in Fig. 5B).

Deletions are formed after initiation of DNA replication. When the ends of transfected linear molecules are joined, the two exons for T antigen become linked, thereby permitting expression of T antigen and SV40 replication. If the deletions that formed at such high efficiency during SVX460 transfection were generated after end joining but before replication, the deleted genomes should be readily detectable among the replicating molecules present in the first cycle of infection, that is, during the first 3 days after transfection. Although the results in Fig. 2 suggest that deleted genomes are not detectable in the first 3 days after transfection, we examined this result in a more sensitive way.

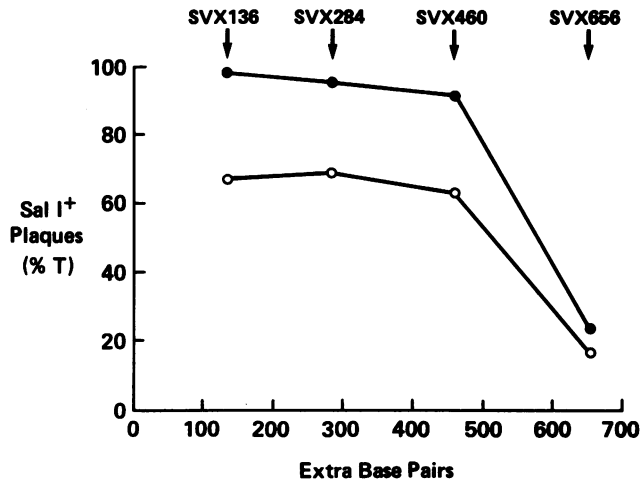


FIG. 6. Frequencies of *SalI*$^+$ plaques generated by oversized SVX genomes. The frequency of *SalI*$^+$ plaques is expressed as a percentage of the total plaques. Filled circles indicate results with circular genomes; open circles indicate results with *SalI*-cleaved linear genomes. The data are from Table 2.

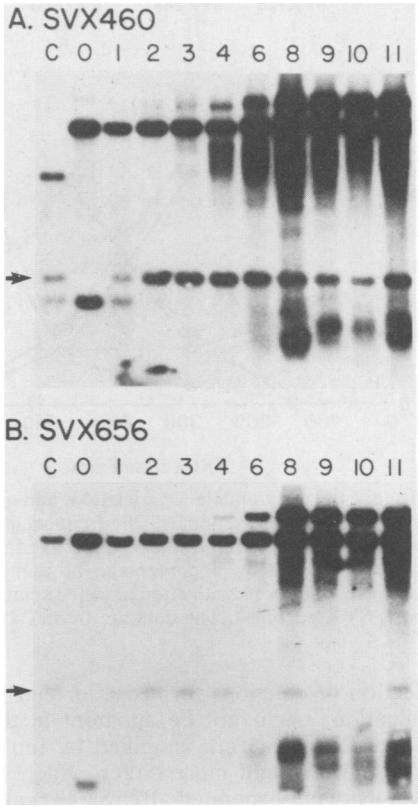


FIG. 7. Deletion formation subsequent to the start of replication. SVX460 and SVX656 genomes linearized by cleavage with *SalI* were transfected onto slightly subconfluent monolayers of CV-1 cells at 10 ng per 60-mm plate. Immediately after transfection (day 0) and at various times thereafter, viral DNA was harvested from two plates, digested with *NdeI*, fractionated by electrophoresis on 1% agarose gels, transferred to Zetabind membranes, hybridized with nick-translated pXB284 DNA, and visualized by autoradiography. Sample volumes were adjusted to give more nearly equal hybridization (days 0 to 1, 30 μl; days 2 to 4, 10 μl; day 6, 3 μl; day 8, 1.5 μl; day 9, 0.5 μl; days 10 to 11, 0.2 μl). The control (C) lane in panel A contained *Bam*HI-linearized SVX460 genomes digested with *NdeI*; the control (C) lane in panel B contained SVX656 circular genomes digested with *NdeI*. In each case, an arrow marks the fragment that would be generated by end joining. The defined fragment at the top of each gels from days 4 to 6 onward represents partial digestion by *NdeI*.

We transfected linear SVX460 and linear SVX656 into CV-1 cells, harvested viral DNA at various times after transfection, and restricted the DNA with *NdeI*, which cleaves SV40 DNA twice. The fragments were separated by electrophoresis on agarose gels and visualized by blot hybridization with nick-translated pXB284 DNA as a probe. Autoradiographs of these analyses are shown in Fig. 7. For both SVX460 and SVX656, the first replicating species was the end join product. The diagnostic restriction fragment generated by end joining is indicated by the arrows in the figure. The first deletion species, which appear as a smear of hybridization around the two *NdeI* fragments, were detectable only after 4 to 6 days.

These results and those in Fig. 2 are consistent with the notion that deletions were formed after SV40 replication was initiated. However, they are not conclusive; a heteroge-

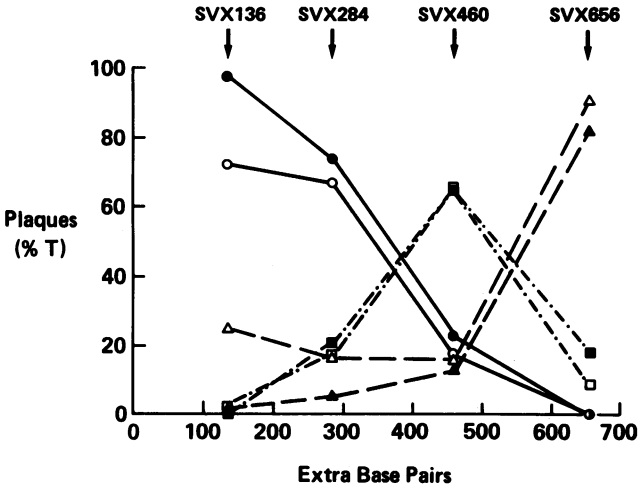


FIG. 8. Frequencies of nondeletion, deletion, and mixed plaques generated by oversized SVX genomes. The frequencies of nondeletion (solid lines), deletion (dashed lines), and mixed plaques (dot-dash lines) are expressed as a percentage of the total number generated during transfections with circular (filled symbols) or linear (open symbols) SVX genomes. The data are from Table 2.

neous collection of deletions formed after end joining, but before replication, might not be apparent in this analysis until after the deletions were amplified by replication and selection. Thus we sought more direct evidence that deletions in SVX460 were generated after SV40 replication was initiated.

Under our standard conditions of plaque assay after DNA transfection, more than 95% of plaques are initiated by single molecules (7, 15; Weiss and Wilson, unpublished data). If the deletions in SVX460 were generated after end joining but before replication, each isolated plaque would be expected to contain a single species of deleted genome. By contrast, if the deletions were generated after replication was initiated, each plaque might contain more than one type of deleted genome.

To check this expectation, the individual plaques that were previously classified as *Sal*I$^+$ or *Sal*I$^-$ were reclassified as nondeletion, deletion, or mixed according to the following criteria. Plaques were classified as nondeletion if they contained all the fragments expected from the circular form of the reference SVX genome and no other fragments. Plaques were classified as deletion if they were missing one or a few contiguous reference fragments and contained one new fragment. Plaques with more complex fragment patterns were classified as mixed. Representative examples of the analysis are shown in Fig. 5, and the results of this classification are summarized in Table 2 and presented graphically in Fig. 8.

Transfection with SVX460, in particular, yielded an extremely high frequency of mixed plaques. To check that these mixed plaques were generated from cells that received single infecting molecules, we repeated the analysis of SVX460 using plaques isolated after transfection with 50-fold-less DNA. The proportion of mixed plaques was the same regardless of DNA concentration, indicating that the mixed plaques were not derived from mixed infections. (The results of transfections at different DNA concentrations were added together in Table 2.) The extremely high frequencies of mixed plaques in SVX460 transfections are a striking confirmation that these plaques were generated after SV40 replication was initiated.

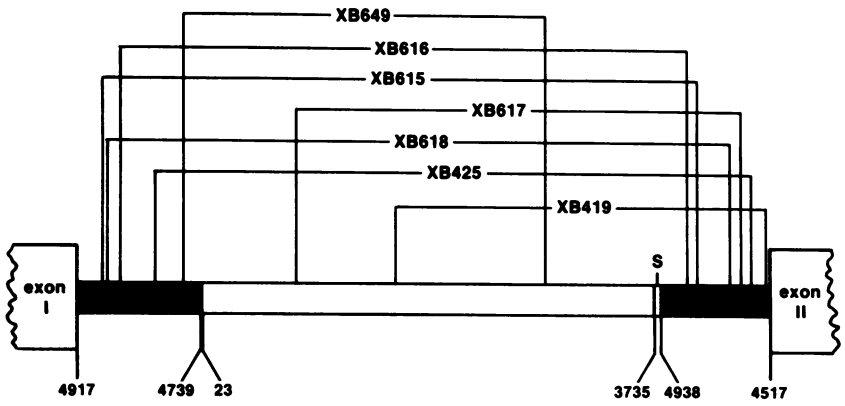
In several ways this analysis of individual plaques confirms the transfection results shown in Fig. 2. SVX136 yielded primarily nondeletion plaques; SVX284 yielded about twice as many nondeletion plaques as other types, consistent with the ratio of full-length to deletion genomes shown in Fig. 2; and SVX656 yielded primarily deletion plaques. Only the finding of apparent nondeletion plaques after transfection of SVX460 requires further comment. As might be expected from the results in Fig. 2, we were unable to clone a nondeleted genome from DNA prepared from four SVX460 nondeletion plaques (data not shown). Given the high frequency of mixed plaques from SVX460 transfections and our inability to clone a full-length genome, it seems likely that the apparent nondeletion plaques actually contained a highly heterogeneous mixture of deleted genomes, which could mimic a nondeletion in our analysis.

Sequences of deletion junctions. The smear of deletion products below both wild-type restriction fragments in Fig. 7 along with the absence of defined bands indicates that the deletions were distributed around the genome. To characterize the deletion junctions, we prepared DNA from seven viable deletions that arose after transfection of *Sal*I-cleaved SVX656 and determined the nucleotide sequences across the junctions. The deletion junctions were all located within the intron, as expected for viable deletions (Fig. 9A). In Fig. 9B the nucleotide sequences around the deletion junctions are shown. The parental sequences present in the deleted genomes are shown in bold type, and the parental sequences that were eliminated in the deletion event are shown in regular type. These junctions are all of the flush junction type (7, 17), with the parental segments joined to one another directly. Homologies exactly at the junctions are indicated by the boxed nucleotides in the figure; they range from 0 to 4 nucleotides and appear quite similar to other junctions we have characterized (7, 17).

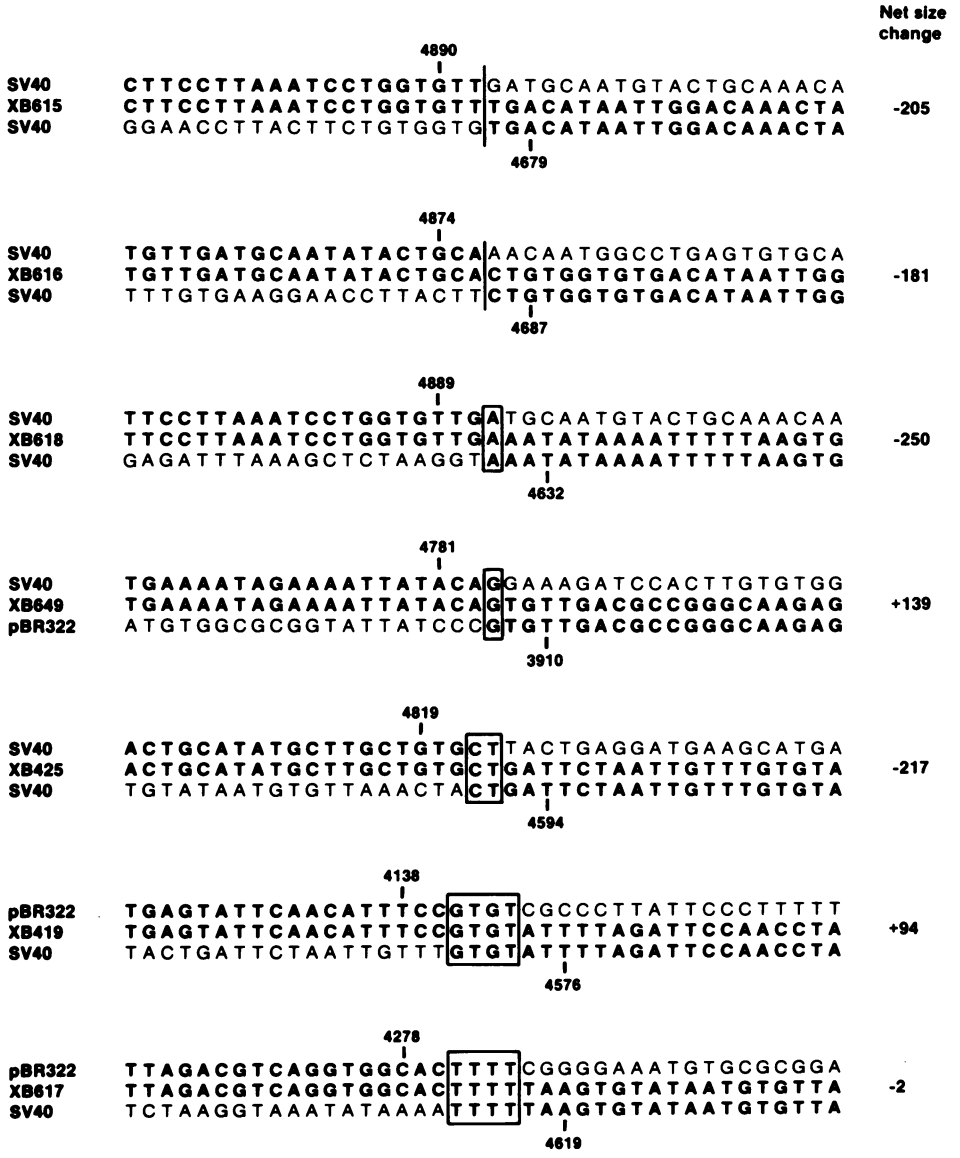
Two arguments suggest that the majority of these junctions arose after end joining. First, deletion XB649 contains the *Sal*I site at which the transfected genomes were originally linearized, indicating that its ends were joined before the deletion event. Second, the most common kind of deletion event during transfection of linear genomes removes sequences only from one end, leaving the other end intact (7). However, none of the sequenced deletions shown

FIG. 9. Deletion junctions in viable progeny derived from transfection with *Sal*I-cleaved linear SVX656 genomes. (A) Distribution of junctions. An expanded diagram of the intron in SVX656 is shown. Solid bars represent SV40 sequences; open bars represent pBR322 sequences. Deletion endpoints are indicated above the intron; SV40 and pBR322 nucleotide numbers are included for reference below the intron (1, 11). (B) Nucleotide sequences around the deletion endpoints. The upper and lower sequences correspond to the parental DNA. Each strand is written 5′ to 3′ in the same sense as the SV40 early mRNA. The nucleotide sequence of each deletion is shown in bold type. The parental sequences that were retained in the deletion event (the left half of the top parental sequence and the right half of the bottom parental sequence) are shown in bold type; the parental sequences that were eliminated in the deletion event are shown in regular type. Vertical bars and boxes show the positions of the junctions. Homologous nucleotides at the junctions are included in the boxes. One nucleotide in each strand is numbered for reference (1, 11). The net size change relative to wild-type SV40 is indicated in base pairs on the right side of the figure.

A


XB649
XB616
XB615
XB617
XB618
XB425
XB419
S
exon I
exon II
4917
4739
23
3735
4938
4517

B


Net size change
4890
SV40 CTTCCTTAAATCCTGGTGTTGATGCAATGTACTGCAAACA
XB615 CTTCCTTAAATCCTGGTGTTTGACATAATTGGACAAACTA -205
SV40 GGAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTA
4679
4874
SV40 TGTTGATGCAATATACTGCAAACAATGGCCTGAGTGTGCA
XB616 TGTTGATGCAATATACTGCACTGTGGTGTGACATAATTGG -181
SV40 TTTGTGAAGGAACCTTACTTCTGTGGTGTGACATAATTGG
4687
4889
SV40 TTCCTTAAATCCTGGTGTTGATGCAATGTACTGCAAACAA
XB618 TTCCTTAAATCCTGGTGTTGAAATATAAAATTTTTAAGTG -250
SV40 GAGATTTAAAGCTCTAAGGTAAATATAAAATTTTTAAGTG
4632
4781
SV40 TGAAAATAGAAAATTATACAGGAAAGATCCACTTGTGTGG
XB649 TGAAAATAGAAAATTATACAGTGTTGACGCCGGGCAAGAG +139
pBR322 ATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAG
3910
4819
SV40 ACTGCATATGCTTGCTGTGCTTACTGAGGATGAAGCATGA
XB425 ACTGCATATGCTTGCTGTGCTGATTCTAATTGTTTGTGTA -217
SV40 TGTATAATGTGTTAAACTACTGATTCTAATTGTTTGTGTA
4594
4138
pBR322 TGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTT
XB419 TGAGTATTCAACATTTCCGTGTATTTTAGATTCCAACCTA +94
SV40 TACTGATTCTAATTGTTTGTGTATTTTAGATTCCAACCTA
4576
4278
pBR322 TTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGA
XB617 TTAGACGTCAGGTGGCACTTTTTAAGTGTATAATGTGTTA -2
SV40 TCTAAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTA
4619

in Fig. 9 were of that type; with the exception of deletion XB649, all the deletions were missing sequences from both ends.

DISCUSSION

Genomes that are just above the length limit for fitting into an SV40 capsid generate deletions with an anomalously high efficiency; one of every two cells infected with such a molecule produces a plaque. The deletion events that are so prevalent in transfected DNA cannot explain such a high frequency of deletion formation; they predict a value at least 10-fold lower (14). To characterize the process responsible for these anomalous deletions, we analyzed transfections involving a series of SV40 genomes that narrowly span the packaging limit. Our analysis of these experiments suggests, somewhat counterintuitively, that the anomalous deletions are actually formed in an otherwise rare process, which is biologically amplified by the packaging properties of the viral capsid.

The packaging limit of the SV40 capsid has been estimated at about 250 extra base pairs by electron microscopic analysis of the lengths of genomes isolated from virus capsids (12). We have reexamined these estimates by measuring the retention of full-length genomes over multiple cycles of infection. Full-length genomes were the majority species at late times after transfections with SVX136 and SVX284, but were undetected at late times after transfections with SVX460 and SVX656 (Fig. 2). These results were confirmed by analysis of genomes present in individual plaques isolated after transfection; the majority of plaques from SVX136 and SVX284 transfections contained nondeleted genomes, whereas few, if any, of the plaques from SVX460 and SVX656 transfections contained nondeleted genomes (Table 2). These data indicate that the effective packaging limit for an SV40 capsid lies between 284 and 460 extra base pairs.

It is probably inappropriate to think of the packaging limit as an exact number of extra base pairs. For example, at late times after transfections of SVX284, which is within the defined packaging limit, deleted genomes accumulated to about 25% of the total (Fig. 2 and Table 2). This accumulation of deletions suggests that smaller genomes are packaged more efficiently than SVX284 and are selected for in multiple rounds of infection. Rather than a precise cutoff, the packaging limit may be better thought of as a steep gradient of decreasing packaging probability. How far this packaging gradient extends is unclear; however, it may include SVX460, since full-length genomes persisted until day 8 after SVX460 transfections, but only until day 4 after SVX656 transfections (Fig. 2).

Although SVX460 and SVX656 are not packaged efficiently, they nevertheless form plaques at higher frequencies (10-fold for SVX460 and two- to threefold for SVX656) than expected based on our measured parameters of breakage and rejoining in transfected DNA (14). In addition, the deletion events underlying these anomalously high infectivities occur after SV40 replication is initiated, whereas deletions in transfected DNA occur before replication (3, 5, 7, 14). The timing of the anomalous deletion events is supported by several observations. The similar infectivities of circles and linears (Fig. 3) and the retention of sequences at the termini of linear genomes (Fig. 4 and 6) indicate that deletions occur after end joining. The appearance of deletions after the first cycle of infection (Fig. 2 and 7) and the high frequency of mixed plaques in SVX460 transfections (Table 2 and Fig. 8) indicate that deletions occur after initiation of SV40 DNA replication.

The timing of deletions after replication and the characteristics of packaging near the limit suggest an explanation for the anomalously high infectivity of slightly oversized genomes. If slightly oversized genomes were packaged and transmitted between cells with low efficiency as suggested above, rare deletion events could be significantly amplified. For example, if the SVX460 genome were packaged at low efficiency, it could replicate not only in the initially infected cell but also in surrounding cells. Multiple rounds of replication would offer this genome increased opportunity to generate deletions that could be packaged more efficiently, thereby allowing it to produce plaques at a higher efficiency than expected if breakage of the input molecules were the only mechanism for deletion. In addition, multiple opportunities for selection of packageable deletions would lead naturally to plaques containing mixtures of different genomes, which for SVX460 accounted for the majority of all plaques. Because the infectivity of SVX656 is lower than that of SVX460, but is still a factor of 2 to 3 above that expected by breakage of input molecules, the SVX656 genome may be packaged with very low efficiency and occasionally transmitted to surrounding cells. The decreased opportunity to generate deletions would lead to fewer mixed plaques, which for SVX656 accounted for only 13% of all plaques.

Variable packaging by an SV40 capsid could be explained in the following two ways: (i) genomes above the effective packaging limit occasionally condense into a smaller volume than usual, thereby permitting their encapsidation; and (ii) capsids occasionally encompass a larger than usual volume, thereby encapsidating a longer genome than usual. More refined measurements will be needed to distinguish between these alternatives.

Implicit in this explanation for high-efficiency plaque formation by oversized genomes is the idea that deletions are generated at low frequency during or after replication. The relatively simple mixtures of genomes in mixed plaques (Fig. 5) and the absence of detectable deletions until 4 days after transfection (Fig. 2 and 7) support this idea. The mechanism by which these deletions were generated is not yet clear, although they may have arisen by breakage of the replicated genomes followed by end joining in a process similar to that responsible for the majority of nonhomologous recombinants in transfected DNA. The characteristics of the deletion junctions in several putative replicative deletions (Fig. 9B) were indistinguishable from the junctions in deletions that were generated before replication began (7, 17), suggesting that the enzymatic machinery responsible for both types of deletions may be similar.

ACKNOWLEDGMENTS

We thank Kathleen Marburger and De Dieu for expert technical assistance. We thank David Roth, Tom Porter, Ursula Weiss, and Al Edwards for valuable suggestions and the Molecular Biology Group in the Department of Biochemistry at Baylor for helpful criticism. We thank Peter Berget in particular for several seminal discussions.

This work was supported by grants from the Public Health Service (CA15743 and GM33405) and the Robert A. Welch Foundation (Q-977).

LITERATURE CITED

1. **Buchman, A. R., L. Burnett, and P. Berg.** 1980. The SV40 nucleotide sequence, p. 799–829. *In* J. Tooze (ed.), DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
2. **Hirt, B.** 1967. Selective extraction of polyoma DNA from

infected mouse cells. J. Mol. Biol. **26**:365–369.
3. **Lebkowski, J. S., R. B. DuBridge, E. A. Antell, K. S. Greisen, and M. P. Calos.** 1984. Transfected DNA is mutated in monkey, mouse, and human cells. Mol. Cell. Biol. **4**:1951–1960.
4. **McCutchen, J., and J. Pagano.** 1968. Enhancement of the infectivity of SV40 DNA with DEAE-dextran. J. Natl. Cancer Inst. **41**:351–357.
5. **Razzaque, A., S. Chakrabarti, S. Joffee, and M. Seidman.** 1984. Mutagenesis of a shuttle vector plasmid after passage in mammalian cells. Proc. Natl. Acad. Sci. USA **80**:3010–3014.
6. **Rigby, P. W. J., M. Diekmann, C. Rhodes, and P. Berg.** 1977. Labelling DNA to high specific activity *in vitro* by nick translation with DNA polymerase I. J. Mol. Biol. **13**:237–251.
7. **Roth, D. B., T. N. Porter, and J. H. Wilson.** 1985. Mechanisms of nonhomologous recombination in mammalian cells. Mol. Cell. Biol. **5**:2599–2607.
8. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
9. **Sharp, P. A., B. Sugden, and J. Sambrook.** 1973. Detection of two restriction endonuclease activities in haemophilus parainfluenzae using analytical agarose-ethidium bromide electrophoresis. Biochemistry **12**:3055–3063.
10. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503–517.
11. **Sutcliffe, J. G.** 1978. Complete nucleotide sequence of Escherichia coli plasmid pBR322. Cold Spring Harbor Symp. Quant. Biol. **43**:77–90.
12. **Tai, H. T., C. A. Smith, P. A. Sharp, and J. Vinograd.** 1972. Sequence heterogeneity in closed simian virus 40 deoxyribonucleic acid. J. Virol. **9**:317–325.
13. **Tooze, J. (ed.).** 1980. DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
14. **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.
15. **Wake, C. T., F. Vernaleone, and J. H. Wilson.** 1985. Topological requirements for homologous recombination among DNA molecules transfected into mammalian cells. Mol. Cell. Biol. **5**:2080–2089.
16. **Wilson, J.** 1978. Interference in SV40 DNA infections: a possible basis for cellular competence. Virology **91**:380–388.
17. **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.
18. **Wilson, J. H., M. DePamphilis, and P. Berg.** 1976. Simian virus 40-permissive cell interactions: selection and characterization of spontaneously arising monkey cells that are resistant to simian virus 40 infection. J. Virol. **20**:391–399.
19. **Wu, R., E. Jay, and R. Roychoudhury.** 1976. Nucleotide sequence analysis of DNA. Methods Cancer Res. **12**:87–176.
20. **Zagursky, R. J., K. Baumeister, N. Lomax, and M. L. Berman.** 1985. Rapid and easy sequencing of large linear double-stranded DNA and supercoiled plasmid DNA. Gene Anal. Techniques **2**:89–94.