MOLECULAR AND CELLULAR BIOLOGY, Oct. 1984, p. 1951–1960
0270-7306/84/101951-10$02.00/0

Vol. 4, No. 10

Transfected DNA Is Mutated in Monkey, Mouse, and Human Cells

JANE S. LEBKOWSKI, ROBERT B. DuBRIDGE, ELIZABETH A. ANTELL, KAY S. GREISEN, AND MICHELE P. CALOS*

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305

Received 19 March 1984/Accepted 26 June 1984

Papovavirus-based shuttle vectors containing the bacterial *lacI* gene were used to show that a mutation frequency in the range of 1% occurs in *lacI* when such vectors are transfected into COS7 and CV-1 simian cells, NIH 3T3, 3T6, L, and C127 mouse cells, and human 293 and HeLa cells. This frequency is approximately four orders of magnitude higher than the spontaneous mutation frequency in either mammalian or bacterial cells. The mutations are predominantly base substitutions and deletions and also include insertions from the mammalian genome. Time course experiments argue that mutagenesis occurs soon after arrival of the DNA into the nucleus. However, replication of the vector is not required since mutations occur even when the vector lacks all viral sequences. The high mutation frequency appears to be the characteristic outcome of transfection of DNA into mammalian cells.

Mammalian cells in culture readily take up DNA. Only a small portion of this transfected DNA reaches the nucleus, and yet a smaller fraction is stably incorporated into the chromosomes in a situation permitting its expression. Although the transfection process is poorly understood, transfection of DNA has become a tool of pivotal importance in the study of the molecular biology of mammalian cells. Application of the technique has facilitated study of topics such as the expression of manipulated genes and the identification of sequences with oncogenic potential. We have sought to use DNA transfection to study mutation in mammalian cells by incorporating an easily scored marker on a transfected vector. Previously, the *lacI* gene of *Escherichia coli* was used as the mutational target on simian virus 40 (SV40)-based vectors that could replicate in both COS7 simian cells and bacteria. We found that, after transfection by the DEAE-dextran technique and 2-day passage in COS7 cells, 1% of the plasmids rescued to *E. coli* contained mutations in *lacI* (3). A similar mutation frequency was detected in related vectors, using the *galK* marker (29). This strikingly high mutation frequency and the facility of the *lacI* system for detailed analysis of the mutations have led us to extend this approach by using different vectors and cell lines, to characterize further the mutational fate of transfected DNA.

In this study, a series of vectors was introduced into a variety of mammalian cell lines to establish the generality of the high mutation frequency. A description of the host DNA that has in some cases become joined to the vectors is also included. The kinetics of mutagenesis were examined by performing time course experiments. To determine whether replication is required for mutagenesis, we introduced molecules lacking a viral origin of replication.

Our observations argue that all transfected DNA, with or without any viral DNA, appears extremely susceptible to mutagenesis in mammalian cells. These findings are compatible with studies that indicate that cotransfected DNA species can become covalently joined, both by homologous recombination and by blunt-end ligation, and may suffer rearrangement (36, 38, 40). Studies of SV40 virus in serial passage (15) and upon integration (2) also reveal rearrangements of the input DNA. The sensitivity of the *lacI* detection system has allowed us to extend this work considerably and to conclude that base substitutions are occurring in the transfected DNA at rates comparable to those of rearrangements. A general picture emerges of a sizable mutagenic effect upon transfection that is probably ubiquitous among mammalian cells. We consider how incoming DNA could be singled out for mutagenesis, as well as the enzymatic activities in the nucleus that could give rise to this phenomenon. We also discuss the practical and biological implications of the high mutation frequency of transfected DNA.

MATERIALS AND METHODS

Plasmid constructions. The construction of pSVi2 and pSVi4 has been described (3). To produce the remaining vectors, the 1.7-kilobase (kb) *Eco*RI fragment from pMC9 (3) was ligated to complete *Eco*RI digests of pMLRIIGΔH3 and pSV2gpt and partial *Eco*RI digests of pJYM, pSV5gpt, and pBRBK100. The ligation mixtures were transformed into CSH35, and red colonies on MacConkey lactose plates, indicating presence of a *lac* operator on the plasmid, were picked (see reference 25). The desired constructions were found among these candidates by examining their restriction patterns. Plasmids were moved to HB101 and purified after chloramphenicol amplification, using a scale up of the alkaline procedure (1) followed by centrifugation through ethidium bromide-cesium chloride gradients.

Detection of I^- mutants. Plasmid DNA was extracted from mammalian cells by the Hirt (13) procedure. As noted in the text, the extracted DNA was frequently treated with *Dpn*I (Boehringer Mannheim). To ensure good recovery of colonies after *Dpn*I treatment, the digests were brought to 100 mM EDTA, extracted with phenol-chloroform-isoamyl alcohol and ether, and ethanol precipitated. The DNA was transformed into MC1061 F'150 kan (see reference 3). In this background plasmids containing wild-type *lacI* will give rise to white colonies on medium containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside, whereas plasmids containing an I^- mutation will give rise to blue colonies. Thus, the proportion of blue to white colonies is a measure of the frequency of I^- mutants in the plasmid population extracted from the mammalian cells. Plasmids from bacterial colonies were purified by the alkaline procedure (1).

Transfections. Mammalian cells were grown in Dulbecco

* Corresponding author.

modified Eagle minimum essential medium supplemented with 10% fetal calf serum. Cells were plated on 60- or 90-mm dishes the day before transfection such that they were approximately 60% confluent at the time of transfection. Transfection was done by the DEAE-dextran procedure (23), using 20 to 100 ng of plasmid DNA per 60-mm plate. Calcium phosphate-mediated transfection was done according to reference 37 with 1 to 5 μg of plasmid DNA per plate without carrier DNA, except that the coprecipitates were left on the cells overnight.

Hybridizations. Genomic DNA was isolated as described by Shih and Weinberg (31) except that lysis was carried out at 56°C and a DNA spooling step (22) was added before phenol extraction. Samples of 9 μg of genomic DNA were digested with *Eco*RI (Boehringer Mannheim), electrophoresed on 0.8% agarose gels, and transferred to Gene Screen Plus (New England Nuclear) according to the protocol provided by the supplier. The transfers were probed with plasmid DNA labeled with ^{32}P by nick translation. Hybridization was performed according to the New England Nuclear protocol, except that 65°C was used and the hybridization buffer consisted of 6× SSC (SSC = 0.15 M NaCl plus 0.015 M sodium citrate), 1% sodium dodecyl sulfate, 5× Denhardt solution, 200 μg of yeast RNA per ml, 100 μg of denatured salmon sperm DNA per ml, and 10^7 cpm of probe (10 to 50 ng; specific activity, $\geq 2 \times 10^8$). Washing was done under high-stringency conditions according to the New England Nuclear protocol at 65°C. Hybridized filters were wrapped in Saran Wrap and exposed to Kodak XAR-5 film for 1 to 3 days.

RESULTS

Simian cells. To study mutagenesis in simian cells, we have used a transient replication protocol that involves placing the *lacI* gene of *E. coli* on vectors that contain the SV40 origin of replication. The vectors also carry the origin of replication of the bacterial plasmid pBR322 and its ampicillin resistance gene. In most of the experiments the plasmids were introduced into COS7 simian cells by the DEAE-dextran technique (for calcium phosphate coprecipitation, see below). The vectors can replicate in these cells because they contain an SV40 origin, and T antigen is provided constitutively from an SV40 provirus with a defective origin of replication (9). The plasmids were allowed to replicate in the mammalian cells for 48 h, extracted by the Hirt (13) procedure, and returned to *E. coli* where mutations in *lacI* were scored as blue colonies on plates containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside.

The plasmids tested in COS7 cells are shown in Fig. 1. We reported *lacI* mutation frequencies of 0.95 and 2.6% for pSVi2 and pSVi4, respectively (3; and extended here). The difference in mutation frequency between the two plasmids, which have the same composition but a different arrangement, is based on the larger nonselected target size for deletions in pSVi4 (see reference 3). The nature of the mutations was determined by preparing plasmid DNA from a large number of I^- mutants and digesting it with *Eco*RI. The size of the *lacI* fragment in the mutant was compared with the 1.7-kb wild-type version. Each mutant was classified as a putative point mutation if the fragment did not differ from wild type, as a deletion if the fragment was shorter, or as an insertion or duplication if it was longer than wild type. The numbers and percentages of mutations falling into each of these categories are shown in Table 1. For pSVi2, where *lacI* is flanked by selected or screened regions, point mutations make up 57% of the I^- events, whereas for pSVi4 deletions predominate (81%) since large deletions can be tolerated.

We have shown previously (3) that the mutations are a product of the plasmids' passage in the mammalian cell. These studies (3) also showed that *lacI* itself was not responsible for the high mutation frequency, since the parent plasmid without *lacI*, pML-RIIG-ChTK (Fig. 1A), has a frequency of rearrangement similar to that of pSVi2 and pSVi4. When randomly chosen white colonies were examined, the rearrangement frequency for the plasmid as a whole was in the range of 10% for each of these vectors. In contrast, a related plasmid, pJYM, containing all of SV40 and lacking the chicken thymidine kinase gene has a 10-fold lower rearrangement rate (3). We therefore asked if specific sequences other than *lacI* could be affecting the rearrangement frequency of the vectors by examining the mutation frequency of a series of SV40-based vectors. The 1.7-kb *Eco*RI *lacI* fragment was added to pJYM in both orientations, producing pJYMia and pJYMib (Fig. 1B). The *Hin*dII fragment containing the chicken thymidine kinase gene was removed from pML-RIIG-ChTK and *lacI* was added, to yield pΔH3ia and pΔH3ib (Fig. 1C). *lacI* was also added to the widely used vector pSV2gpt (27), producing pSV2gptia and pSV2gptib (Fig. 1D). The resulting mutation frequencies and the nature of the mutations are shown in Table 1. The mutation frequencies range from 0.3 to 2.5% and generally correlate with the amount of DNA contiguous with *lacI* that can be deleted and still give a surviving I^- plasmid that would be scored in our detection system.

To determine whether special features of the COS7 cell line were responsible for the mutagenic effect, pJYMia and pJYMib were also transfected into CV-1 simian cells, which is the line from which COS7 was derived by SV40 transformation. pJYM derivatives can replicate in CV-1 cells since these vectors produce their own T antigen. The frequency and nature of the mutations obtained in CV-1 cells closely resemble that observed in COS7 (Table 1). We also compared the effect of using calcium phosphate coprecipitation instead of DEAE-dextran as the method of DNA introduction. In these experiments, larger amounts of DNA were used, generally 1 to 5 μg per plate, without carrier DNA, compared to 20 ng for the DEAE-dextran experiments. The amount of residual nonreplicated input DNA that persists throughout the experiment is not significant when 20 ng of DNA is transfected. However, the amount of nonreplicated input DNA is large when several micrograms of DNA are introduced. Therefore, when plasmid DNA was extracted from the calcium phosphate-treated cells after replication, the DNA was treated with *Dpn*I to destroy residual input DNA. This restriction enzyme recognizes the sequence GATC only when the adenine residue is methylated (16). This modification is made in bacterial but not mammalian cells. Therefore, treatment with *Dpn*I will leave intact only the DNA that replicated in mammalian cells (28). To monitor the completeness of *Dpn*I digestion, the following control was included for each set of samples treated with *Dpn*I. Bacterially grown plasmid DNA was added to a Hirt extract prepared from mock-transfected cells. This sample was digested with *Dpn*I in parallel with the samples containing replicated DNA. Tranformation of the control sample into *E. coli* resulted in few or no bacterial colonies. We found that the mutation frequency and the nature of the mutations detected are similar whether the DNA is introduced with calcium phosphate or DEAE-dextran (data not shown).

Mouse and human cells. To learn whether the high mutation frequency could be generalized to other species and tissue types or was confined to monkey kidney cells, we

tested vectors that would replicate in mouse and human cells. *lacI* was added to pSV5gpt (27), producing pPyia and pPyib (Fig. 2a). These plasmids contain the early region from polyoma virus and can replicate in mouse cells. One or both of these vectors were transfected into mouse L, NIH 3T3, 3T6, and C127 cells and into BHK-21 Syrian hamster cells, where the vectors also replicate. Transfection of 3T3 and 3T6 cells was done with DEAE-dextran, whereas the calcium phosphate technique was used to transfect the BHK cells. Both methods were used to introduce DNA into C127 and L cells. Replication was allowed to proceed for 2 to 6 days before the plasmid DNA was harvested by the Hirt procedure. In all cases, the extracted DNA was treated with *Dpn*I to eliminate input DNA. The completeness of *Dpn*I digestion was verified as described above. The frequency of I^- mutations in the plasmid pools extracted from mouse cells is shown in Table 2, as well as the nature of the I^- mutations.

The study was extended to human cells, using the SV40-based vectors pJYMia and pJYMib (Fig. 1B) and the BK-based vectors pBKia and pBKib (Fig. 2B). We anticipated that the pJYMi vectors would replicate in human cells since human cells are semipermissive for SV40 replication and the vectors code their own T antigen. BK is a human papovavirus closely related to SV40 and has been reported to replicate in a number of human cell lines (35). Neither the pJYMi nor the pBKi vectors replicated in WI-38 human embryonic lung cells or human whole fetal cells. However, both vectors replicate well in 293, an adenovirus 5-trans-

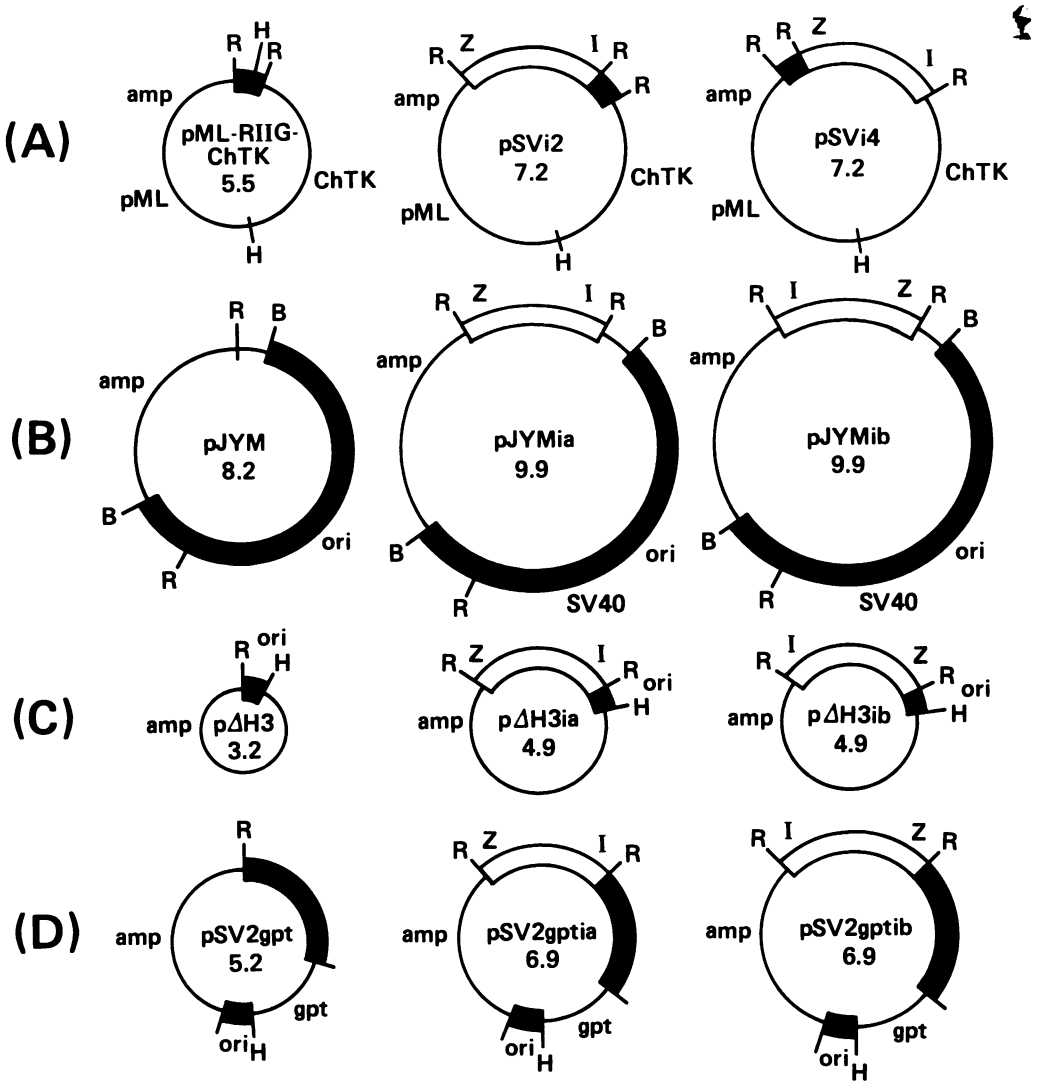

FIG. 1. SV40-based plasmids. All plasmids are drawn to scale. The heavy line represents SV40 DNA. The open line denotes *lac* sequences and the single line represents pML and other vector sequences. The 1.7-kb *lac* fragment contains the wild-type *lacI* gene, as well as the beginning of *lacZ*. The *lacZ* portion is necessary to give a blue colony in our assay (3). (A) pML-RIIG-ChTK 4, 21) includes the chicken thymidine kinase gene and a 312-base pair fragment containing the SV40 origin of replication (heavy line). pSVi2 and pSVi4 contain the 1.7-kb *lacI* fragment inserted on opposite sides of the SV40 origin (3). (B) pJYM contains all of SV40 inserted in the *Bam*HI site of pML (21). pJYMia and pJYMib contain the 1.7-kb *lac* fragment inserted into the *Eco*RI site of pML in both orientations. (C) pΔH3 was made by deleting from pML-RIIG-ChTK the *Hin*dIII fragment which contains the chicken thymidine kinase gene and a portion of the SV40 fragment (not including the origin of replication). The 1.7-kb *lac* fragment was added to the *Eco*RI site in both orientations. (D) pSV2gpt (27) contains sequences derived from SV40 including the origin of replication, the *E. coli gpt* gene, and part of pBR322. The 1.7-kb *lac* fragment was added in both orientations. B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI.

TABLE 1. Mutations derived from monkey cells[a]

Transfection into:	Plasmid	No. of I^-/ total no. of colonies	Frequency of I^- (%)	No. examined	% Point mutation (no.)	% Deletion (no.)	% Insertion/ duplication (no.)
COS7 cells	pSVi2	305/30,940	1.0	225	57 (129)	38 (86)	4 (10)
	pSVi4	1,134/45,934	2.5	96	13 (12)	81 (78)	6 (6)
	pJYMia	445/87,100	0.5	205	28 (58)	70 (143)	2 (4)
	pJYMib	143/30,243	0.5	79	49 (39)	44 (35)	6 (5)
	pΔH3ia	83/29,473	0.3	40	53 (21)	45 (18)	3 (1)
	pΔH3ib	254/36,051	0.7	45	51 (23)	38 (17)	11 (5)
	pSV2gptia	196/14,405	1.4	62	13 (8)	81 (50)	6 (4)
	pSV2gptib	40/8,168	0.5	28	25 (7)	68 (19)	7 (2)
CV-1 cells	pJYMia	259/33,468	0.8	49	18 (9)	82 (40)	0 (0)
	pJYMib	105/29,796	0.4	10	40 (4)	60 (6)	0 (0)

[a] The numbers for each plasmid represent pooling of at least several independent dishes of cells in this and the following tables. Cells were transfected by the DEAE-dextran method with 20 ng of DNA. The number examined in this and the following tables refers to the number of I^- colonies for which plasmid DNA was prepared and digested with *Eco*RI to determine the nature of the mutation.

formed derivative of human embryonic kidney cells. Apparently the adenovirus 5-transforming functions or other aspects of 293 permit replication of the vectors, a phenomenon we are investigating. The pJYMi vectors also replicated at moderate levels in HeLa cells, a human epithelial cervical carcinoma line. Plasmid DNA was introduced into human cells with the calcium phosphate technique and allowed to replicate for several days. The Hirt extracts were treated with *Dpn*I to eliminate input DNA. Table 3 shows the I^- frequencies for the recovered plasmids and the nature of the I^- mutations.

Kinetics of mutagenesis. Time course experiments were undertaken to determine when mutagenesis occurs and whether the mutation frequency is established immediately or increases throughout the replication period. To detect changes in mutation frequency with the greatest sensitivity, we used pSVi4 for these experiments. It has the highest mutation frequency, 2.5%. Twelve 60-mm plates of COS7 cells were transfected with 50 ng of pSVi4 DNA, using the DEAE-dextran technique. The plates were harvested at intervals of 0 to 53 h and plasmid DNA was extracted. Half of the sample was treated with *Dpn*I to remove input DNA before being returned to *E. coli*. The other half was not *Dpn*I treated and thus contains both input and replicated plasmid DNA. *Dpn*I-resistant colonies were absent at 0, 6, and 9 h and first appeared at 20 h (43 colonies). The first I^- colony

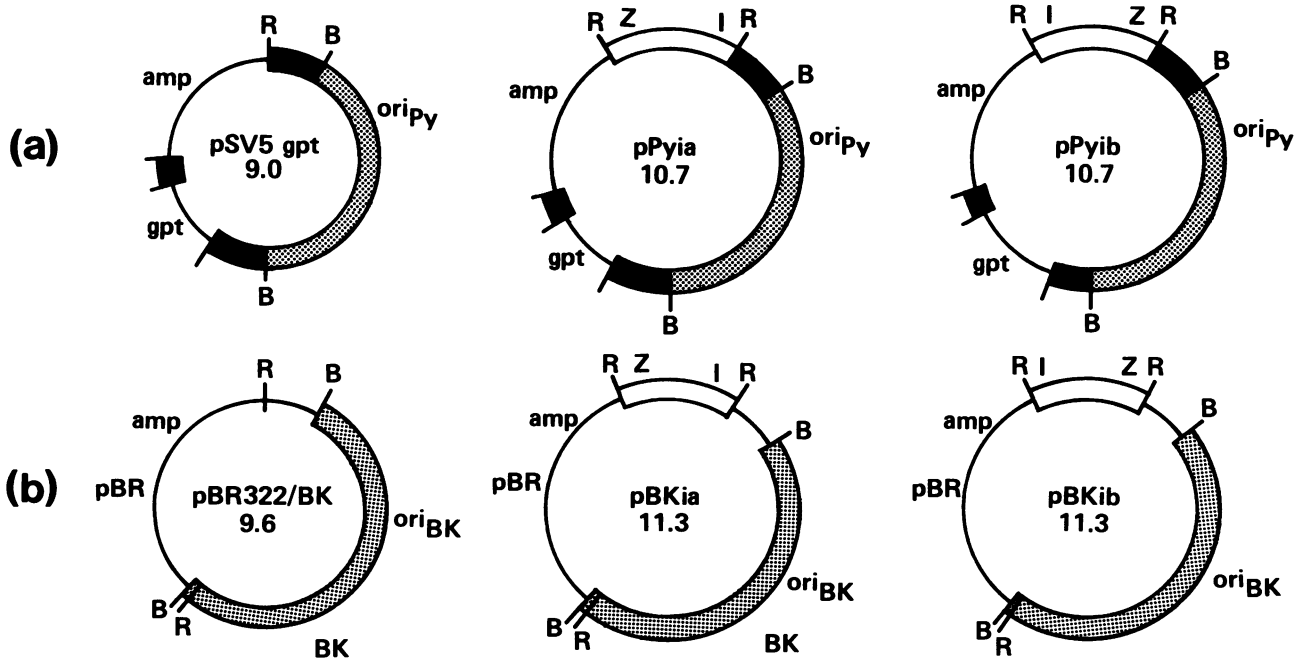

FIG. 2. Polyoma- and BK-based vectors. The stippled line represents polyoma sequences in (a) and BK sequences in (b). The solid line denotes SV40 sequences. The open line represents the 1.7-kb *lacI*-containing fragment, and the single line represents other procaryotic vector sequences. (a) pSV5gpt is pSV2gpt (Fig. 1D) with a portion of polyoma encompassing the replication origin and early region inserted at the *Bam*HI site. The *lac* fragment was inserted into the *Eco*RI site in both orientations to give pPyia and pPyib. (b) pBR322/BK contains all of BK virus inserted into the *Bam*HI site of pBR322. The *lac* fragment was inserted at the *Eco*RI site of pBR322 in both orientations to give pBKia and pBKib.

TABLE 2. Mutations derived from rodent cells[a]

Transfection into:	Plasmid	No. of I^-/total no. of colonies	Frequency of I^- (%)	No. examined	% Point mutation (no.)	% Rearrangement (no.)
L cells	pPyia	8/1,579	0.5	8	0 (0)	100 (8)
	pPyib	102/9,940	1.0	26	8 (2)	92 (24)
3T3 cells	pPyia	12/563	2.1	12	0 (0)	100 (12)
	pPyib	9/1,001	0.9	9	44 (4)	55 (5)
3T6 cells	pPyia	60/6,913	0.9	25	12 (3)	88 (22)
C127 cells	pPyib	17/2,054	0.8	23[b]	17 (4)	83 (19)
BHK-21 cells	pPyib	7/507	1.4	7	0 (0)	100 (7)

[a] Transfections were done with DEAE-dextran or calcium phosphate, as noted in the text. Mutations were classified as rearrangements if the *lacI* *Eco*RI fragment was other than wild type in size.
[b] Additional I^- colonies were obtained from experiments in which the I^- frequency was not determined.

appeared as one of 60 colonies at 25 h. Thereafter, the colony number generally increased with time, reaching over 3,000 at 53 h. However, the I^- frequency did not increase with time, but rather varied from plate to plate with a range of 1.0 to 3.7% (mean, ca. 2.6%). Thus, the data do not substantiate a progressive increase in mutation frequency incurred through replication. The results are consistent with the mutation frequency being achieved early, without significant alteration during the entire replication period. Without *Dpn*I treatment, I^- mutations also first appear at 25 h, even though over 4,000 colonies were observed before this time.

The *Dpn*I-treated time course indicates that detectable plasmid replication begins 20 to 25 h after transfection. Because replication of the vectors is presumably confined to the nucleus, we infer that the molecules are in the nucleus at 20 to 25 h. Since I^- mutations first appear at 25 h, the molecules are evidently in the nucleus when detectable mutagenesis occurs. Time course experiments cannot eliminate the argument that DNA damage occurs before arrival in the nucleus and that this damage can only be processed to a form recoverable as mutations in the nucleus.

Time course experiments with pSVi4 were repeated twice, using 100 ng of DNA per plate, and analyzed with and without *Dpn*I treatment. Similar experiments were also performed with the pSVi2 vector, using either calcium phosphate or DEAE-dextran as the mode of DNA introduction, and analyzed with and without *Dpn*I treatment. In all of these cases, replication of the vectors as signalled by the appearance of *Dpn*I-resistant colonies commenced at approximately 20 to 30 h after transfection. The first I^- mutations appeared during the same period. In no case did the advent of I^- mutations precede the start of the replication period. Similarly, in no case did the *Dpn*I-treated samples show an increase in mutation frequency with time. Instead they varied randomly about a mean value typical for the vector used. Over 200 I^- mutations gathered throughout the time courses were examined to determine the nature of the mutations. Both point mutations and rearrangements were found throughout the period in which I^- mutations were detected.

A nonreplicating vector. Having determined that mutagenesis occurs at about the time that replication of the vectors commences in the mammalian cell, it was pertinent to inquire whether replication itself generates the mutations or whether the nuclear environment alone is sufficient to produce mutations. By eliminating the *Eco*RI fragment containing the SV40 origin of replication from pSVi4, we created a molecule, pSVi ori$^-$, which lacks a eucaryotic origin of replication. We verified that this plasmid cannot replicate in mammalian cells (see below). pSVi ori$^-$ is similar to pSVi4 in that I^- mutations can be generated by a point mutation or rearrangement within *lacI* or by a deletion extending from *lacI* into the neighboring nonselected chicken thymidine kinase gene. Its mutation frequency might thus be expected to be similar to that of pSVi4 at 2.5%.

Eight plates of COS7 cells were transfected with 100 ng of pSVi ori$^-$ DNA, using the DEAE-dextran method. The plates were harvested at approximately 6-h intervals from 6 to 48 h. The extracted DNA was not treated with *Dpn*I since, lacking the ability to replicate in mammalian cells, all of the DNA remains sensitive to *Dpn*I. Mutations were absent at the 6-, 12-, 18-, and 24-h time points, but appeared at 29 h

TABLE 3. Mutations derived from human cells[a]

Cells	Plasmid	No. of I^-/total no. of colonies	Frequency of I^- (%)	No. examined	% Point mutation (no.)	% Rearrangement (no.)
293	pJYMib	175/365,252	0.05	91	68 (62)	32 (29)
	pBKia	15/10,861	0.14	14	29 (4)	71 (10)
	pBKib	13/27,390	0.05	13	85 (11)	15 (2)
HeLa	pJYMia	2/1,995	0.10	11[b]	73 (8)	27 (3)
	pJYMib	2/2,460	0.08	2	0 (0)	100 (2)

[a] Transfections were done with calcium phosphate. Mutations were classified as rearrangements if the *lacI* *Eco*RI fragment was other than wild type in size.
[b] Additional I^- colonies were obtained from experiments in which the I^- frequency was not determined.

TABLE 4. Mutations incurred in a nonreplicating plasmid, pSVi ori$^-$ [a]

Transfection into:	Time after transfection (h)	No. of I^-/total no. of colonies	Frequency of I^- (%)	No. examined	No. of point mutations	No. of rearrangements
COS7 cells	24–36	22/54,160	0.04	22	15	7
CV-1 cells	24–29	10/50,067	0.02	10	5	5
NIH 3T3 cells	24–44	15/64,576	0.02	13	11	2

[a] A 100-ng amount of pSVi ori$^-$ DNA per 60-mm dish was transfected by the DEAE-dextran method. The totals represent pooling of data obtained at several time points within the range given, including dishes from which no i$^-$ mutants were obtained.

and later. Therefore, the nonreplicating plasmid is mutated, and mutagenesis follows a time course paralleling that obtained with the replicating vectors. Since a replicating plasmid yields *Dpn*I-resistant colonies at 29 h and can therefore be assumed to be in the nucleus at this time, we infer that the nonreplicating plasmid is also located in the nucleus by 29 h. This observation suggests that mutation again occurs in the nucleus, but replication per se is not required.

Transfection of pSVi ori$^-$ into COS7 cells was repeated twice in the same manner, focusing on time points between 24 and 33 h. These data are summarized in Table 4. The I^- frequency for pSVi ori$^-$ in COS7 was 0.04%. The 60-fold discrepancy between this frequency and the 2.5% frequency for pSVi4 can be largely or completely accounted for if the frequency is being artificially depressed by large amounts of input DNA, most of which is not in the nucleus and therefore not subject to mutation. The amount of residual input DNA present at a given time is dependent on the amount of DNA originally transfected. We can demonstrate that, when 100 ng of DNA is used, a substantial amount of input DNA remains at 22 h. By using pSVi4, the relative fractions of input versus replicated DNA can be determined by assaying the sample with and without *Dpn*I. The numbers of colonies present, as well as the percentage of I^- colonies, was examined at early times when 100 ng of pSVi4 was transfected in the same manner as pSVi ori$^-$. At 22 h, the pSVi4 sample not treated with *Dpn*I yielded 1,412 colonies, whereas the *Dpn*I-treated pSVi4 sample gave only 86 colonies. Therefore, >90% of the DNA present at 22 h is input. This fraction drops progressively over time for pSVi4 as replicated plasmid accumulates. However, for pSVi ori$^-$ the high fraction of input would remain. The mutation frequency of the pSVI4 22-h sample not treated with *Dpn*I is only 0.28% (4 of 1,412), reflecting the high fraction of input DNA. The frequency rises to approximately 2.5% as the proportion of replicating DNA, with its I^- molecules, increases. The I^- frequency of 0.28% for pSVi4 at 22 h is comparable to the value obtained for pSVi ori$^-$, 0.04%. The higher value for pSVi4 probably reflects the fact that the plasmid was already replicating by 22 h, as evidenced by the presence of colonies from *Dpn*I-resistant DNA. Such colonies actually were detected as early as 17 h in this experiment.

Since the mutation frequency of pSVi ori$^-$ in COS7 cells is low, it was important to verify that we were not merely detecting the spontaneous mutation frequency of the plasmid in *E. coli*. Therefore, pSVi ori$^-$ DNA was extracted from COS7 cells immediately after transfection and returned to *E. coli*. The mutation frequency was determined to be 0.00062% (6 of 958,800) or 6.2×10^{-6}. This frequency is similar to the spontaneous mutation rate reported for *lacI* carried on an F' *lac pro* episome of 2.5×10^{-6} (26). The mutation frequency we detected for pSVi ori$^-$ in COS7 cells is more than 50-fold higher than the spontaneous background and can be safely ascribed to alterations requiring the mammalian cells. To rule out involvement of T-antigen in mutagenesis, we introduced pSVi ori$^-$, which contains no viral information, into CV-1 simian cells and NIH 3T3 mouse cells which contain no papovaviral sequences. Since the mutation frequency for pSVi ori$^-$ in CV-1 and NIH 3T3 is similar to the frequency measured in COS7 cells (Table 4), the mutagenesis cannot be attributed to the presence of viral T antigen.

pSVi ori$^-$ has no viral replication. To verify that it indeed lacks the ability to replicate in mammalian cells, samples of the DNA extracted from CV-1 and NIH 3T3 cells were divided into three portions. Each portion was treated with *Dpn*I, *Mbo*I, or no enzyme and was transformed into *E. coli*. *Mbo*I represents the converse of *Dpn*I in that it can cut GATC sequences only if the adenine is not methylated (8). Therefore, only DNA that did replicate in the mammalian cell will be sensitive to *Mbo*I, whereas these same molecules would be resistant to *Dpn*I. The pSVi ori$^-$ extracts were completely sensitive to *Dpn*I, indicating that they did not replicate in the mammalian cells. The extracts were resistant to *Mbo*I, and these *Mbo*I-treated extracts yielded i$^-$ colonies. Thus, pSVi ori$^-$ DNA retained its bacterial modification pattern and several of these nonreplicated molecules acquired mutations. The mutations that arose in pSVi ori$^-$ were examined, as reported in Table 4. Both point mutations and rearrangements were detected.

Insertions. We wished to characterize the DNA insertions in more detail, particularly to address the possibility that some of the insertions might represent mammalian transposable elements. A collection of eight I^- mutations that produced a net increase in the size of the plasmid (Fig. 3) and were not duplications were examined further by restriction and hybridization analysis. The mutations occurred in pSVi2 and pJYMib during their replication in COS7 cells and were therefore expected to contain simian DNA. Accordingly, each plasmid was labeled by nick translation and hybridized to a filter containing genomic DNA from COS7 cells digested with *Eco*RI (Fig. 4) and also to filters containing DNA derived from *E. coli* MC1061. All of the plasmids showed hybridization to COS7 DNA and none to *E. coli* DNA. Nick-translated parental pSVi2 or pJYMib was hybridized to *Eco*RI-digested COS7 DNA in parallel with each hybridization of insert-containing plasmids. No hybridization of the parental plasmids to COS7 DNA was detected, verifying that the hybridization was due to the insert DNA. Six of the plasmids hybridized intensely to the monkey DNA, indicating that the inserts contained repeated sequences. Two of the plasmids, 12-9 and 17-7 (Fig. 4D), hybridized to the *Eco*RI-digested COS7 DNA in a ladder pattern characteristic of the α-satellite, a repeated sequence found in approximately 10,000 copies per haploid cell, both dispersed and in tandem arrays (5, 32). That these inserts indeed contained α sequences was confirmed by their hybridization to an α-specific probe (data not shown).

A further two inserts, 12-15 and 22-61, contain very highly repeated sequences which we expected might be members of

the *Alu* repeat family (Fig. 4C), present in approximately 300,000 copies per haploid genome (14). This prediction was confirmed by the specific hybridization of these plasmids to the BLUR-8 *Alu*-containing probe (data not shown). Two plasmids, 10-15 and 22-1, hybridized to a repeat family present in a lower copy number (Fig. 4B and E). They did not hybridize to the α or *Alu* probe, or to a probe specific for the *Kpn* family, a dispersed and heterogeneous repeat family present to the extent of approximately 4,000 to 50,000 copies in haploid primate genomes (11). Therefore, plasmids 10-15 and 22-1 may represent previously uncharacterized repeat families. These repeat families are also present in the human genome, as indicated by hybridization of plasmids 10-15 and 22-1 to human DNA derived from the 293 cell line. However, hybridization to 293 DNA is much weaker than that to COS7 DNA (data not shown). The remaining plasmids, 9-25 and 22-29, do not hybridize to repeated sequences and presumably represent single or low-copy sequences derived from the COS7 genome (Fig. 4A and F).

DISCUSSION

We have conclusively demonstrated that DNA transfected into mammalian cells suffers an extraordinarily high mutation frequency. Although all cell lines tested show elevated mutation frequencies, there are differences between the species examined. The mutation frequency is in the neighborhood of 1% for a variety of simian and murine cell lines, but only about 0.1% for two human cell lines. When *lacI* is flanked by selected or screened sequences in the vector, the mutations are approximately equally divided between point mutations and rearrangements in the primate cells tested, whereas rearrangements predominate in the rodent cells.

Sequence analysis through genetic techniques shows that most of the mutations classified as point mutations are base substitutions, and these substitutions occur throughout the gene (J. H. Miller, J. S. Lebkowski, K. S. Greisen, and M. P. Calos, submitted for publication). Restriction analysis of a set of the deletions shows that the endpoints are randomly distributed throughout the gene (3). The insertions also appear random with respect to their target site within the gene and with respect to the host DNA inserted. Most (six of eight) of the inserts examined contain repeat sequences. This result is not unexpected given the sizes of the inserts characterized and the distribution of repeats in the genome. Furthermore, since the wide variety of vector constructions tested in this study all mutate in a similar way, it is unlikely that a specific sequence on the vector DNA is triggering mutagenesis.

Sampling the transfected DNA at different times after introduction into recipient cells establishes two points: first, mutations can only begin to be recovered after approximately 24 h. This is also the approximate time that the first *Dpn*I-resistant plasmids can be recovered from the cells, that is, the first plasmids replicated by the mammalian cell. This correspondence suggests that the mutations occur in the nucleus, soon after arrival there. Second, the frequency of I^- mutants does not show a progressive rise over time, but rather varies randomly about a mean typical for the plasmid being examined. This result is in accord with the data of Razzaque et al. (29) which indicate a constant mutation frequency for a *galK* vector at 24, 48, 72, and 96 h after transfection. These findings fortify the conclusion that the mutations are introduced early, rather than continually, during replication.

Having confined the time of mutagenesis to the period soon after arrival of the vectors in the nucleus, the possibilities remained that either mutagenesis occurs during the early rounds of replication or alterations occur in the DNA independent of replication. Experiments with a vector lacking a viral origin of replication, and consequently unable to replicate in mammalian cells, demonstrate that all classes of mutation can be obtained without overall replication of the vector, though we cannot rule out the possibility that some mutations might also form after replication has begun. If the DNA is mutated in the nucleus, the lesions could be caused by attack by nuclear enzymes, or they could represent nuclear processing of damage preexisting in the DNA. Since most transfected DNA is degraded by the cell (17), we hypothesized that the mutations could represent nuclear processing of DNA partially degraded in the cytoplasm, possibly in lysosomes. Direct microinjection experiments of

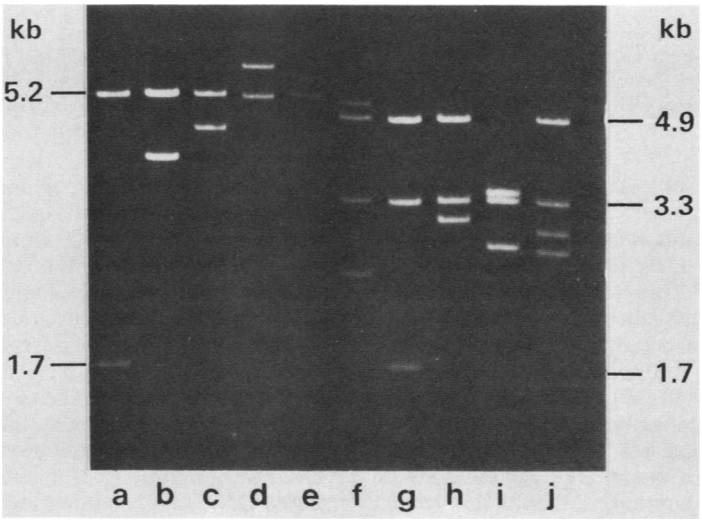

FIG. 3. Plasmids containing insertions in *lacI*. Ethidium bromide-stained 0.8% agarose gel showing the *Eco*RI digestion patterns for the parental plasmids (pSVi2 and pJYMib) and their insertion-containing derivatives. Lane a, pSVi2 and its derivatives; b, 9-25; c, 10-15; d, 12-15; e, 12-9; g, pJYMib and its derivatives; f, 17-7; h, 22-1; i, 22-29; j, 22-61.

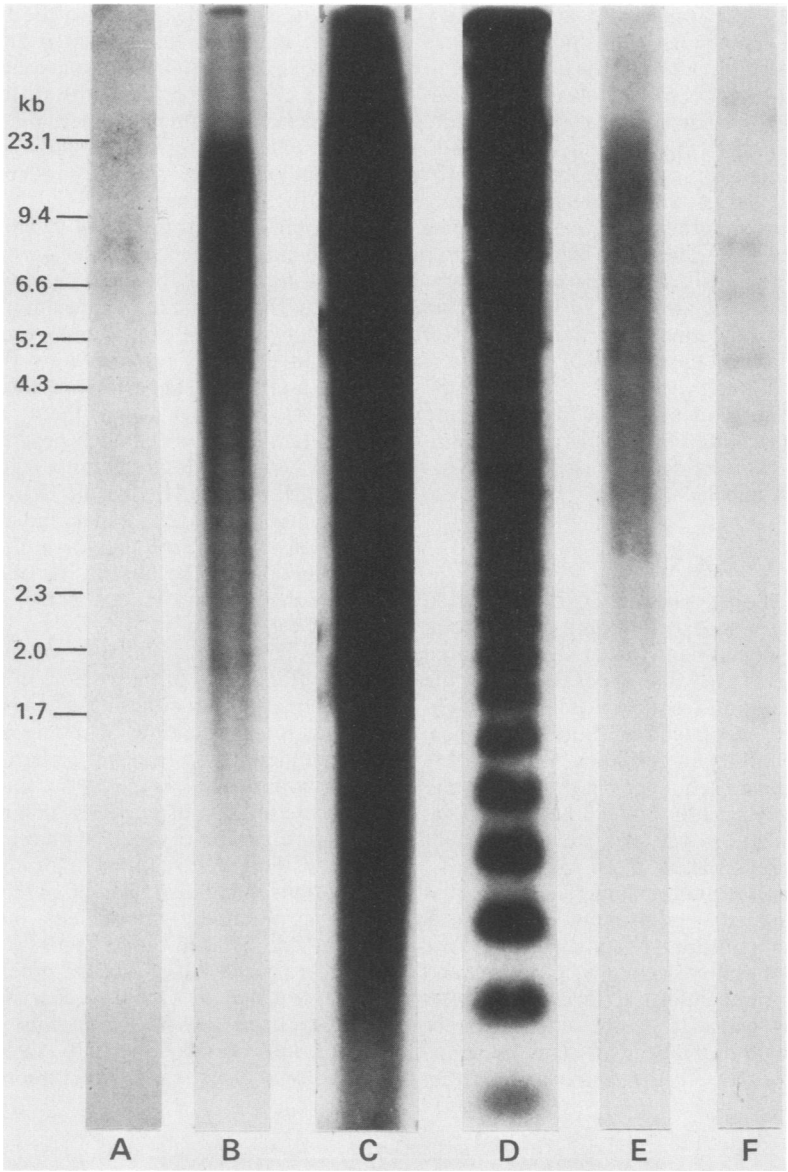

FIG. 4. Hybridization of insertions to simian DNA. Southern analysis, using the ^{32}P-labeled insertion plasmids as hybridization probes against parallel transfers of *Eco*RI-digested COS7 cell DNA. The probes used were nick-translated plasmid DNA from: a, 9-25; b, 10-15; c, 12-15; d, 12-9; e, 22-1; f, 22-29. When either of the ^{32}P-labeled parental plasmids (pSVi2 and pJYMib) was used as a probe, no hybridization was detected. The hybridized filters were exposed to XAR-5 film for 12 h in the case of lanes C and D and 24 h in the case of lanes A, B, E, and F.

the vector DNA into the nuclei of COS7 cells were undertaken to test this hypothesis. Preliminary results with pSVi4 and pJYMib indicate that, at least for rearrangements, microinjected DNA is mutated at a frequency similar to that observed for transfected DNA (M. R. Capecchi, J. S. Lebkowski, and M. P. Calos, unpublished data).

A variety of endonucleases and exonucleases reside in the mammalian nucleus (18, 19). Abundant ligase activity has also been documented (6, 24, 38, 39). If newly transfected DNA is subject to attack by these enzymes, deletions and insertions could readily be generated. As for point mutations, we know that most of them are single-base pair substitutions (Miller et al., submitted). All available routes to base substitution seem to require polymerase action. Such a polymerase would have to be extensively error prone to account for the observed mutation frequency. We note that, when assayed in vitro, mammalian polymerases have no associated proofreading capacity and exhibit error rates that are compatible with the mutation frequencies we observe (20). Similarly, high mutation frequencies also occur during the somatic generation of diversity in the immune system (34). Further experiments are planned to elucidate the mechanism of formation of base substitutions in our system.

The mutagenesis we have detected in *lacI* appears to be representative of the fate of any transfected DNA. There exists extensive evidence that heterogeneous transfected DNAs efficiently become covalently joined by ligation (38, 40). Furthermore, rearrangements are detectable in transfected DNA that has integrated into the chromosome (2, 2a, 36). "Recombinations at non-homologous sequences," which we would describe as deletions, have been described in transfected vectors that can replicate (33, 39). Rearrange-

ments in viral DNA introduced as virions have long been studied as defective viruses. These structures, as extensively analyzed in SV40 (12, 15), have the form of the rearrangements observed in transfected DNA and may arise in the same way. Point mutations would not readily have been detected in previous experiments. This important class of mutations has only come to light with shuttle vector systems. However, point mutations presumably affect all transfected DNA. This may be the explanation for the detection of an activated *ras* oncogene selected after transfection of NIH 3T3 cells with the wild-type c-*ras* version of the gene (30). The rare activated transfectant was found to contain a *ras* gene which had acquired a G:C to A:T transition in the 12th codon. The likelihood that alterations are randomly introduced into transfected DNA at appreciable rates must obviously be considered in interpreting experiments involving transfection.

The extent to which the mammalian nucleus manipulates incoming DNA is impressive. What distinguishing feature does the newly transfected DNA possess which makes it liable to an extraordinary mutation frequency? We have not eliminated the argument that the procaryotic provenance of the DNA, evidenced for example in its pattern of methylation, marks it for attack. However, an obvious feature of the transfected DNA that sets it apart from the chromosomal DNA is that the incoming DNA enters the cell naked and free of chromatin. This uncomplexed state may be an excellent substrate for degradative enzymes and may mimic their intended targets, such as replication forks and recombination junctures. We do not know when the transfected DNA becomes complexed with chromosomal proteins. However, the demonstration that mutagenesis takes place early after arrival of the DNA in the nucleus may indicate that the mutations occur before a normal chromatin conformation is attained. High rates of recombination (6) and gene expression (7, 10) have also been reported for newly transfected DNA. All of these effects may be mediated by some "exposed" aspect of the incoming DNA. The reactivity of transfected DNA, as opposed to the relative inertness of chromosomal DNA, may reflect the absence of various nuclear proteins that would normally limit activity.

ACKNOWLEDGMENTS

We are grateful to Suzanne Clancy for constructing pSV2gptia, pSV2gptib, and pMLΔH3ib. We thank M. Botchan for helpful discussions and for providing pΔH3 and pBRBK100 DNA and C127 cells, V. Oi for pSV5gpt DNA and CV-1, BHK 21, and NIH 3T3 cells, C. Hsu for 293 cells, R. Kornberg and D. Goldfarb for HeLa cells, R. Cassin for 3T6 cells, M. Zolan and C. A. Smith for alpha probe, P. Kavathas for BLUR 8 probe and L cells, M. Singer for *Kpn* probe, and G. Nolan and H.-J. S. Huang for contributing data to Table 3.

J.S.L. was supported by Public Health Service Cancer Biology Training grant CA-09302 from the National Cancer Institute. R.B.D. and E.A.A. were supported by Public Health Service predoctoral training grants. This work was funded by Public Health Service grant CA-33056 awarded to M.P.C. by the National Cancer Institute and by grant 83-E-101 from the Chicago Community Trust/Searle Scholars Program.

ADDENDUM IN PROOF

Recent experiments of Razzaque et al. (A. Razzaque, S. Chakrabarti, S. Joffee, and M. Seidman, Mol. Cell. Biol. 4:435–441, 1984) examined rearrangements in transfected shuttle vectors. They conclude that the majority of deletion and insertion mutants are generated before replication of the vector, in agreement with our results.

LITERATURE CITED

1. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
2. **Botchan, M., J. Stringer, T. Mitchison, and J. Sambrook.** 1980. Integration and excision of SV40 DNA from the chromosome of a transformed cell. Cell **20**:143–152.

2a. **Bullock, P., W. Forrester, and M. R. Botchan.** 1984. DNA sequence studies of simian virus 40 chromosomal excision and integration in rat cells. J. Mol. Biol. **174**:55–84.

3. **Calos, M. P., J. S. Lebkowski, and M. R. Botchan.** 1983. High mutation frequency in DNA transfected into mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **80**:3015–3019.
4. **Conrad, S. E., C.-P. Liu, and M. R. Botchan.** 1982. A fragment spanning the SV40 replication origin is the only DNA sequence required in cis for viral excision. Science **218**:1223–1225.
5. **Fittler, F.** 1977. Analysis of the α-satellite DNA from African green monkey cells by restriction nucleases. Eur. J. Biochem. **74**:343–352.
6. **Folger, K. R., E. A. Wang, G. Wahl, and M. R. Capecchi.** 1982. Patterns of integration of DNA microinjected into cultured mammalian cells: evidence for homologous recombination between injected plasmid DNA molecules. Mol. Cell. Biol. **2**:1372–1387.
7. **Gaynor, R. B., and A. J. Berk.** 1983. *Cis*-acting induction of adenovirus transcription. Cell **33**:683–693.
8. **Gelinas, R. E., P. A. Myers, and R. J. Roberts.** 1977. Two sequence-specific endonucleases from *Moraxella bovis*. J. Mol. Biol. **144**:169–179.
9. **Gluzman, Y.** 1981. SV40-transformed simian cells support the replication of early SV40 mutants. Cell **23**:175–182.
10. **Green, M. R., R. Treisman, and T. Maniatis.** 1983. Transcriptional activation of cloned human β-globin genes by viral immediate-early gene products. Cell **35**:137–148.
11. **Grimaldi, G., and M. F. Singer.** 1983. Members of the *KpnI* family of long interspersed repeated sequences join and interrupt α-satellite in the monkey genome. Nucleic Acids Res. **11**:321–337.
12. **Gutai, M. W., and D. Nathans.** 1978. Evolutionary variants of SV40: cellular DNA sequences and sequences at recombinant joints of substituted variants. J. Mol. Biol. **126**:275–288.
13. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cultures. J. Mol. Biol. **26**:365–369.
14. **Jelinek, W. R., T. P. Toomey, L. Leinwand, C. H. Duncan, P. A. Biro, P. V. Chondary, S. M. Weissman, C. M. Rubin, C. M. Honik, P. L. Deininger, and C. W. Schmid.** 1980. Ubiquitous interspersed repeated sequences in mammalian genomes. Proc. Natl. Acad. Sci. U.S.A. **77**:1398–1402.
15. **Kelly, T. J., and D. Nathans.** 1977. The genome of simian virus 40. Adv. Virus Res. **21**:85–173.
16. **Lacks, S., and B. Greenberg.** 1975. A deoxyribonuclease of *Diplococcus pneumoniae* specific for methylated DNA. J. Biol. Chem. **250**:4060–4066.
17. **Lebkowski, J., M. Botchan, and M. Calos.** 1983. Spontaneous mutation of DNA transfected into mammalian cells, p. 603–614. *In* E. C. Freidberg and B. R. Bridges (ed.), Cellular responses to DNA damage. UCLA Symposia on Molecular and Cellular Biology, new ser., vol. 11. Alan R. Liss, Inc., New York.
18. **Lindahl, T.** 1982. DNA repair enzymes. Annu. Rev. Biochem. **51**:61–87.
19. **Linn, S. M., and R. J. Roberts (ed.).** 1982. Nucleases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. **Loeb, L. A., and T. A. Kunkel.** 1982. Fidelity of DNA synthesis. Annu. Rev. Biochem. **51**:429–457.
21. **Lusky, M., and M. Botchan.** 1981. Inhibition of SV40 replication in simian cells by specific pBR322 DNA sequences. Nature (London) **293**:79–81.
22. **Marmur, J.** 1961. A procedure for the isolation of DNA from microorganisms. J. Mol. Biol. **3**:208–218.
23. **McCutchan, J. H., and J. S. Pagano.** 1968. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethyl aminoethyl-dextran. J. Natl. Cancer Inst. **41**:351–356.
24. **Miller, C. K., and H. M. Temin.** 1983. High-efficiency ligation

and recombination of DNA fragments by vertebrate cells. Science **220**:606–609.
25. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. **Miller, J. H., D. Ganem, P. Lu, and A. Schmitz.** 1977. Genetic studies of the *lac* repressor. I. Correlation of mutational sites with specific amino acid residues: construction of a colinear gene-protein map. J. Mol. Biol. **109**:275–301.
27. **Mulligan, R. C., and P. Berg.** 1980. Expression of a bacterial gene in mammalian cells. Science **209**:1422–1427.
28. **Peden, K. W. C., Y. M. Pipas, S. Pearson-White, and D. Nathans.** 1980. Isolation of mutants of an animal virus in bacteria. Science **203**:1392–1396.
29. **Razzaque, A., H. Mizusawa, and M. M. Seidman.** 1983. Rearrangement and mutagenesis of a shuttle vector plasmid after passage in mammalian cells. Proc. Natl. Acad. Sci. U.S.A. **80**:3010–3014.
30. **Santos, E., E. P. Reddy, S. Pulciani, R. J. Feldman, and M. Barbacid.** 1983. Spontaneous activation of a human proto oncogene. Proc. Natl. Acad. Sci. U.S.A. **80**:4679–4683.
31. **Shih, C., and R. A. Weinberg.** 1982. Isolation of a transforming sequence from a human bladder carcinoma cell line. Cell **29**:161–169.
32. **Singer, D. S.** 1979. Arrangement of a highly repeated DNA sequence in the genome and chromatin of the African green monkey. J. Biol. Chem. **254**:5506–5514.
33. **Subramani, S., and P. Berg.** 1983. Homologous and nonhomologous recombination in monkey cells. Mol. Cell. Biol. **3**:1040–1052.
34. **Tonegawa, S.** 1983. Somatic generation of antibody diversity. Nature (London) **302**:575–581.
35. **Tooze, J.** 1981. DNA tumor viruses: molecular biology of tumor viruses, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
36. **Wigler, M., A. Pellicer, S. Silverstein, and R. Axel.** 1978. Biochemical transfer of single-copy eucaryotic genes using total cellular DNA as donor. Cell **14**:725–731.
37. **Wigler, M., S. Silverstein, L.-S. Lee, A. Pellicer, Y.-C. Cheng, and R. Axel.** 1977. Transfer of purified herpes virus thymidine kinase gene to cultured mouse cells. Cell **11**:223–232.
38. **Wigler, M., R. Sweet, G. K. Sim, B. Wold, A. Pellicer, E. Lacy, T. Maniatis, S. Silverstein, and R. Axel.** 1979. Transformation of mammalian cells with genes from procaryotes and eukaryotes. Cell **16**:777–785.
39. **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.
40. **Winocour, E., and I. Keshet.** 1980. Indiscriminate recombination in simian virus 40-infected monkey cells. Proc. Natl. Acad. Sci. U.S.A. **77**:4861–4865.
41. **Witkin, E.** 1976. Ultraviolet mutagenesis and inducible repair in *Escherichia coli*. Bacteriol. Rev. **40**:869–907.