MOLECULAR AND CELLULAR BIOLOGY, Jan. 1987, p. 379–387
0270-7306/87/010379-09$02.00/0

Vol. 7, No. 1

Analysis of Mutation in Human Cells by Using an Epstein-Barr Virus Shuttle System

ROBERT B. DuBRIDGE,[1] PETER TANG,[1] HAN CHAO HSIA,[2] PHAIK-MOOI LEONG,[2] JEFFREY H. MILLER,[2] AND MICHELE P. CALOS[1]*

Department of Genetics, Stanford University School of Medicine, Stanford, California 94305,[1] and Department of Biology, University of California, Los Angeles, California 90024[2]

Received 30 July 1986/Accepted 9 October 1986

We developed highly sensitive shuttle vector systems for detection of mutations formed in human cells using autonomously replicating derivatives of Epstein-Barr virus (EBV). EBV vectors carrying the bacterial *lacI* gene as the target for mutation were established in human cells and later returned to *Escherichia coli* for rapid detection and analysis of *lacI* mutations. The majority of the clonal cell lines created by establishment of the *lacI*-EBV vector show spontaneous LacI$^-$ frequencies of less than 10^{-5} and are suitable for studies of induced mutation. The ability to isolate clonal lines represents a major advantage of the EBV vectors over transiently replicating shuttle vectors (such as those derived from simian virus 40) for the study of mutation. The DNA sequence changes were determined for 61 *lacI* mutations induced by exposure of one of the cell lines to *N*-nitroso-*N*-methylurea. A total of 33 of 34 *lacI* nonsense mutations and 26 of 27 missense mutations involve G · C to A · T transitions. These data provide support for the mutational theory of cancer.

A venerable goal of the science of genetics is to understand the nature of heritable change. Interest in such mutational change stems from the fundamental importance of the process for generating genetic diversity, as well as the concern inspired by the role of mutation in initiating genetic disease and cancer. Molecular biology can be applied to develop a better understanding of mutagenesis. In particular, recombinant DNA techniques are useful for dissecting the process in complex eucaryotic cells. Shuttle vectors have been constructed which allow the introduction of well-defined bacterial genes as targets for mutation in mammalian cells. These autonomously replicating vectors can be retrieved in bacteria for rapid detection and analysis of mutations, bypassing many of the limitations of genetic analysis in mammalian cells.

The shuttle vector approach was first demonstrated in simian cells, using vectors based on simian virus 40 (SV40). These experiments revealed that the vectors acquired mutations in their bacterial target genes at a frequency of approximately 1%, resulting simply from brief passage of the vectors through the mammalian cells (3, 27). The bulk of the mutations seemed to be due to damage incurred in vector DNA early after its entry into the mammalian cells, rather than during replication. Even nonreplicating molecules were subject to the effect, and the mutation frequency did not increase during the replication period (16).

The high spontaneous mutation frequency of DNA transfected into mammalian cells presented an obstacle to studies of rarer classes of mutation with shuttle vector systems. This obstacle has been met on the one hand by finding mammalian cell backgrounds, e.g., human 293 cells, in which the frequency of transfection-associated mutations is somewhat lower (16). In these cells mutations can be induced above background by using mutagens such as UV light and ethyl methanesulfonate, and the mutations made by these agents have been analyzed (18, 19). Another approach has been to introduce high enough levels of damage into the DNA with the mutagenizing agent before transfection that the level of mutation resulting from this damage will exceed that occurring spontaneously in simian cells (13).

However, the sensitivity of the SV40 shuttle systems is limited by the high frequency of mutation associated with transfection. In addition, SV40 vectors replicate poorly in most human cells (17). A possible solution to these problems was raised by the isolation of autonomously replicating minireplicons derived from the human herpesvirus Epstein-Barr virus (EBV) by Sugden and colleagues (37, 41, 42).

If most of the mutations incurred by transiently replicating vectors are early events targeted to intracellular damage inflicted during transfection, then stably replicating vectors such as those derived from EBV could offer a solution to the problem. The initial establishment of stable vectors via transfection would be expected to lead to some mutation of the incoming DNA. However, since the mutation frequency is about 1% or lower per thousand base pairs (bp), it should be relatively easy to obtain cell lines which carry only wild-type vectors. If no appreciable mutagenesis occurs during vector replication, the mutation frequency of vector DNA purified from such cell lines should remain very low. In this case, stable vector systems would allow one to clone molecules which have safely passed the transfection-mutation barrier. This procedure is not possible in cells transfected with transiently replicating plasmids because vector replication prevents long-term survival of the cells.

The experiments described here show that EBV vectors do display the expected behavior of stably replicating vectors. Lines with exceedingly low mutation frequencies can readily be isolated. Our data support the interpretation that the high mutation frequency observed with transiently replicating shuttle vectors was primarily due to transfection and was not an inherent property of vector replication. Thus, the use of vectors whose continued replication is compatible with cell survival allows cells containing mutation-free populations of vectors to be cloned. This strategy solves the problem of the high mutation frequency associated with transfection.

* Corresponding author.

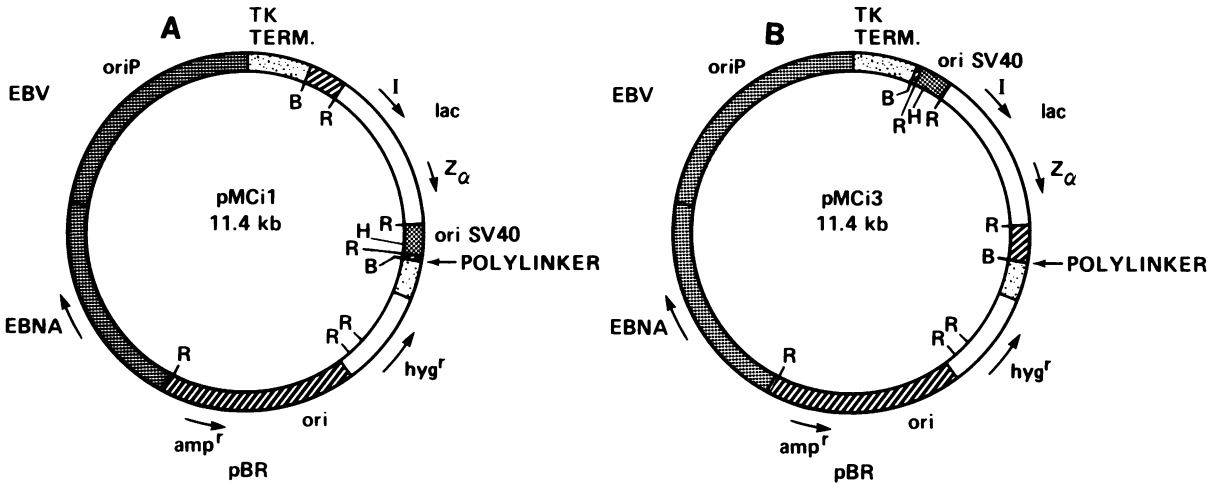

FIG. 1. Structures of pMCi1 and pMCi3. Plasmid constructions are described in the text. Filled portions represent viral (EBV or SV40) sequences, shaded portions are pBR322 sequences, open sections are bacterial genes (*lacI* or Hyg^r), and dotted sections are herpes simplex virus thymidine kinase (TK) control regions. B, *Bam*HI; H, *Hin*dIII; R, *Eco*RI. The other part of the pUC12 polylinker is at the first *Bam*HI site counterclockwise from the site marked polylinker. ori, Origin.

The low spontaneous background of the EBV shuttle systems facilitates isolation of induced mutations. We demonstrate this feature here by mutagenizing a human cell line carrying a *lacI*-EBV vector with the alkylating agent *N*-nitroso-*N*-methylurea (NMU). An increase in mutation frequency above background induced by ethyl nitrosourea has also been demonstrated by Drinkwater and Klinedinst (7) using a related EBV vector system. We collected 225 NMU-induced *lacI* mutations and analyzed them using genetic and biochemical techniques. The DNA sequence changes involved in the mutations suggest the mechanism of NMU mutagenesis.

NMU is of special interest because it is also a potent carcinogen. The mutational theory of cancer would postulate that the activity of NMU as a carcinogen is due to mutations it generates in specific genes involved in oncogenesis. Barbacid and colleagues (38, 43) induced mammary tumors in rats with NMU and isolated activated H-*ras* genes from the tumors. Of 48 tested, 48 differed from the *ras* genes of nontumor tissues by a G · C to A · T transition affecting codon 12. We find that the G · C to A · T transition accounts for 97% of the mutations induced by NMU in mammalian cells. These data support the interpretation favored by Barbacid et al. to explain the role of NMU in carcinogenesis.

MATERIALS AND METHODS

Plasmids. pMCi1 and pMCi3 (Fig. 1) are derivatives of plasmid p220.2, generously provided by B. Sugden. p220.2 is an 8.9-kilobase (kb) plasmid that contains the EBV *oriP* and EBNA-1 sequences, the pBR322 ampicillin resistance gene and origin of replication, the hygromycin resistance gene under the control of promoter and termination sequences from the herpes simplex virus type 1 thymidine kinase gene, and a polylinker site. The EBV sequences include the EBV nuclear antigen 1 (EBNA-1) gene from EBV coordinates 107,567 to 110,176 and the *oriP* region from EBV coordinates 7,333 to 9,516. The EBNA-1 sequences were inserted between the *Hin*dIII site of pBR322 and the *Sph*I site of *oriP* in the plasmid pHEBo, which has been described previously (37). In addition, p220.2 contains the polylinker from pUC12 (*Sma*I-*Hae*III fragment) inserted into the *Nar*I site in the herpes simplex virus thymidine kinase termination sequences. Fragments (2.3 kb) containing *lacI* and the SV40 origin of replication were inserted into the *Bam*HI site of the polylinker in p220.2 to generate pMCi1 and pMCi3. These fragments were derived from plasmids made by inserting the 1.7-kb *lacI* fragment from pMC9 (24) into pBOP (8). The 1.7-kb *lacI* *Eco*RI fragment from pMC9 was ligated to full-length *Eco*RI partial digestion products of pBOP. Inserts of the 1.7-kb *lac* fragment were obtained in both orientations on either side of the 311-bp *Eco*RI fragment of pBOP containing the SV40 origin. Each of these pBOP-*lacI* plasmids could be cut with *Xho*II to produce a 2.3-kb fragment containing the *lac* and SV40 origin fragments. For pMCi1 the 2.3-kb fragment used extends from the *Xho*II site in the β-lactamase gene of pBR322 (bp 4010), through the *lac* 1.7-kb and SV40 0.3-kb *Eco*RI fragments to the *Bam*HI (*Xho*II) site of pBOP. This site is in the position of the *Hin*dIII site of pBR322, which was converted to a *Bam*HI site during the construction of pBOP. This fragment was inserted into the polylinker of p220.2 at the unique *Bam*HI site. The orientations of the *lac* and SV40 fragments are as shown in Fig. 1A. pMCi3 was made by inserting a corresponding 2.3-kb fragment from a pBOP-*lacI* derivative into the *Bam*HI site in the polylinker of p220.2. In this case a derivative of pBOP was used such that the *Xho*II fragment encompasses the *Bam*HI (formerly *Hin*dIII)-*Eco*RI portion of pBR322, followed by the 0.3-kb SV40 origin fragment, the 1.7-kb *lacI* fragment, and the *Eco*RI to *Xho*II (bp 4010) portion of pBR322. The orientations of the *lac* and SV40 *Eco*RI fragments are as shown in Fig. 1B. Thus, pMCi1 and pMCi3 differ in the location of the SV40 origin on one side or the other of the *lac* fragment.

Tissue culture and mutagenesis. Late-passage 293 human embryonic kidney cells (11) were grown in Dulbecco modified Eagle medium plus 10% fetal calf serum, penicillin, and streptomycin in a 10% CO_2 incubator. The cells were transfected with 5 μg of pMCi1 and pMCi3 DNA per 100-mm plate by the calcium phosphate coprecipitation technique (40). Selection was initiated on day 3 after transfection by addition of 200 μg of hygromycin B (Calbiochem-Behring,

La Jolla, Calif.) per ml to the growth medium. This concentration of hygromycin was maintained in the growth medium throughout all the experiments with lines stably carrying EBV vectors. Mutagenesis was performed by dissolving NMU (ICN Biomedicals, K & K Laboratories, Inc., Plainview, N.Y.) in water to make a 10-mg/ml stock solution. This solution was immediately added to the growth medium of cells at 30 to 40% confluence.

Plasmid return to bacteria. Vector DNA was extracted from human cells by the Hirt procedure (14). Hirt extracts were treated with proteinase K, phenol, phenol-chloroform-iosamyl alcohol, and ethanol precipitation. DNA was transformed into frozen competent MC1060 F'150 Kan *recA* cells. The transformation efficiency ranged from 5×10^5 to 2×10^6 colonies per μg of DNA. Cells were plated and LacI$^-$ colonies were detected as described previously (3). A Biotran II automatic colony counter (New Brunswick Scientific Co., Inc., Edison, N.J.) was used to count the colonies.

Mutation analysis. Plasmid DNA was prepared from each LacI$^-$ colony by the small-scale alkaline lysis procedure and examined on a 1% agarose gel after digestion with *Eco*RI (21). Plasmid DNA from the mutants was also introduced into strain GM1 by transformation, and homogenotized blue colonies, representing transfer of the *lacI* mutation to the F' *lac pro* episome (9), were picked. The episomes carrying the *lacI* mutations were gridded and mated into strains carrying nonsense suppressors or *lacI* deletions by replica mating techniques. Identification and characterization of nonsense mutations, as well as detailed genetic mapping of the mutations, were done as described previously (4, 5, 31, 32). For biochemical sequence analysis, the GM1 cells carrying the mutations were infected with an f1 derivative carrying the wild-type *lacI* gene and the α-encoding portion of the *lacZ* gene with an amber mutation at coding position 17 (S. K. Whoriskey and J. H. Miller, unpublished data). Bacteriophage produced from this line were introduced into a *lacI*$^-$ *lacZ* α-acceptor strain, and blue colonies were picked. The *lacI* mutations carried by the phage were sequenced by the dideoxy technique, using appropriate primers based on the location of the mutations as determined by genetic mapping.

RESULTS

Cell lines. Plasmids pMCi1 and pMCi3 (Fig. 1) were constructed. These vectors are derived from p220.2 (gift of Bill Sugden), which contains *oriP*, the origin for EBV latent replication; the *trans*-acting gene coding for EBNA-1, the EBV nuclear antigen, which is required for activation of the origin; a gene conferring resistance to hygromycin for selection in mammalian cells; and portions of pBR322 for selection and replication in *Escherichia coli*. Such vectors are carried stably at copy numbers of approximately 1 to 100 per cell in a wide range of human cells as long as selection pressure for hygromycin is maintained (41, 42). A fragment containing *lacI*, the beginning of *lacZ*, and the SV40 origin of replication was inserted into p220.2 to create pMCi1 and pMCi3.

The plasmids were introduced by calcium phosphate coprecipitation into 293 cells, a line of human embryonic kidney cells transformed by the early region of adenovirus type 5 (11). Hygromycin selection was applied, resulting in the formation of numerous hygromycin-resistant colonies. Two plates transfected with each of the plasmids were allowed to grow to confluence, so that by 9 days after transfection several thousand individual colonies fused to form a population. Each of these four populations was propagated by 1:50 dilution weekly. Populations derived from the two pMCi1 transfection plates were designated 1.3 and 1.4, and those derived from the plates receiving pMCi3 were designated 3.3 and 3.4.

TABLE 1. Human 293 cell lines carrying pMCi1 and pMCi3[a]

Clones	Total colonies	No. of plates	Colonies/plate	No. LacI$^-$	Frequency LacI$^-$
pMCi1					
1.11	154,914	77	2,012	1	6.4×10^{-6}
1.14	12,877	9	1,431	0	
1.15	2,250	6	375	0	
1.16	1,839	6	307	0	
1.21	26,510	12	2,209	0	
1.22	3,352	6	559	0	
1.24	1,151	3	384	0	
1.25	3,477	6	580	39	1.1×10^{-2}
1.26	14,965	6	2,494	27	1.8×10^{-3}
pMCi3					
3.11	6,751	6	1,125	0	
3.12	10,704	9	1,189	0	
3.13	13,811	12	1,151	0	
3.16	1,459	6	243	0	
3.21	2,747	6	458	0	
3.23	12,340	12	1,028	0	
3.24	14,413	9	1,601	0	
3.28	41,961	17	2,468	0	
3.29	11,126	6	1,854	0	
Populations					
pMCi1					
1.3	18,157	8	2,270	3	1.6×10^{-4}
1.4	27,397	10	2,740	9	3.3×10^{-4}
pMCi3					
3.3	30,904	8	3,863	12	3.8×10^{-4}
3.4	24,115	7	3,445	13	5.3×10^{-4}

[a] The number of bacterial colonies recovered from the designated number of 100-mm plates of 293 cells is given. The number of LacI$^-$ colonies and the LacI$^-$ frequency for those lines which yielded mutants are noted. Data are reported for individual clones and for pooled populations.

To obtain clonal lines, two transfection plates derived from each plasmid were split 1:10,000 on day 9, giving rise to four groups of plates designated 1.1 and 1.2 for pMCi1 and 3.2 for pMCi3. On day 26 after the original transfection, colonies were picked from each of the dilution sets into 24-well plates and propagated into mass culture.

Spontaneous mutation frequency. Confluent 100-mm plates of cells derived from the four populations and from nine individual clones for pMCi1 and nine for pMCi3 were harvested. Vector DNA was purified from the cells by the Hirt procedure (14), transformed into *E. coli* MC1061 F'150 Kan *recA*, and plated on medium containing 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-gal). The numbers of bacterial colonies recovered were counted, and LacI$^-$ colonies were scored as blue colonies. LacI$^+$ colonies are white on X-gal in this system (see reference 3). The number of colonies recovered from the clones and the populations, as well as the LacI$^-$ frequencies, are reported in Table 1. Each of the populations showed a relatively high frequency of LacI$^-$ mutants, ranging from 1.6×10^{-4} to 5.3×10^{-4}. These values are comparable to the mutant frequency of 3.5×10^{-4} obtained with the SV40 vector pJYMib after transfection and passage for 3 days in 293 cells (18).

In contrast, the majority of the clones were free of mutation. No mutations were found during initial screening

TABLE 2. Mutation frequencies of transiently replicated or transformed DNA

Plasmid	No. of colonies	No. LacI$^-$	Frequency LacI$^-$
Transient replication in 293/*tsA1609*neo[a]			
pMCi1	17,312	166	9.6×10^{-3}
pMCi3	11,122	15	1.3×10^{-3}
pJYMib	18,566	21	1.1×10^{-3}
Transformation into *E. coli* MC1061 F'150 Kan *recA*[b]			
pMCi1	1,373,700	31	2.3×10^{-5}
pMCi3	1,439,100	21	1.5×10^{-5}
pJYMib[c]	1,918,776	31	1.6×10^{-5}

[a] The number of bacterial colonies recovered from plasmid DNA transfected and replicated in human 293/*tsA1609*neo cells is given, along with the number and frequency of LacI$^-$ mutants.

[b] The number of bacterial colonies obtained after transformation of plasmid DNA directly into *E. coli* and the number and frequency of LacI$^-$ mutant colonies are given.

[c] From reference 18.

for 16 of 18 clones examined. Clone 1.11 was chosen for further study. Of more than 150,000 bacterial colonies formed by transformation of DNA extracted from this line, only one mutation was observed. Therefore, the spontaneous background for *lacI* carried on an EBV vector in 1.11 is below 10^{-5}. Since the mutation frequency for *lacI* plasmids transformed directly into the recipient *E. coli* is approximately 1×10^{-5} to 2×10^{-5} (see Table 2), it is possible that the one mutation obtained from line 1.11 actually formed in the bacterial recipient rather than the human cells. We noticed no increase in mutation frequency of line 1.11 during 9 months of continuous passage. Thus, the line does not appear to harbor LacI$^-$ plasmids nor to generate them at any appreciable frequency.

As Table 1 indicates, there was a variation in the number of colonies recovered from the different clones. The number of colonies recovered per plate of cells for a given line was fairly constant, but the number between clones varied by approximately 1 order of magnitude. The mechanism controlling copy number of EBV and its plasmid derivatives is not presently understood. On the basis of quantification of Southern blots of DNA extracted from 1.11, we estimate that 1.11 carries approximately 10 copies of pMCi1 per cell (L. Lam and M. Calos, unpublished data).

To directly test the mutation frequencies of pMCi1 and pMCi3 under transient replication conditions, we introduced the plasmids into cell line 293/*tsA1609*neo. This line is a derivative of 293 into which the gene for the temperature-sensitive SV40 T-antigen mutant *tsA1609* (28) has been inserted. This line produces replication-competent T antigen in large amounts at 37°C (R. DuBridge and M. Calos, unpublished data). The pMCi1 and pMCi3 plasmids, which carry the SV40 origin of replication but not T antigen, undergo replication driven by the SV40 origin in 293/*tsA1609*neo. The SV40-driven replication greatly exceeds that driven by the EBV origin in a transient replication situation. pMCi1, pMCi3, and pJYMib were transfected into 293/*tsA1609*neo cells, allowed to replicate for 3 days, and then harvested by Hirt extraction. Plasmid DNA was treated with *Dpn*I to remove unreplicated DNA (see reference 16) and returned to MC1061 F'150 Kan *recA*. The mutation frequencies detected are reported in Table 2. The frequency for pMCi3 resembled that obtained for pJYMib. In both these plasmids, *lacI* is flanked by selected or screened sequences, precluding large deletions. The frequency for pMCi1, in which *lacI* is unflanked on one side, was higher. The mutation frequencies for pMCi1 and pMCi3 after transient replication are high and resemble the frequencies previously obtained with other SV40 vectors.

NMU mutagenesis. The low spontaneous mutation frequency of line 1.11 and the relatively high number of bacterial colonies obtained from plasmid DNA extracted from it led us to use this line for studies of induced mutation. The alkylating agent NMU, a potent mutagen and carcinogen (22, 29, 35), was chosen for study. A dose-response curve was determined to examine the effect of increasing NMU concentration on LacI$^-$ frequency. NMU was dissolved in water and immediately added to the growth medium of cultures of 1.11 cells at approximately 30% confluence. Growth medium containing the mutagen, which has a short half-life in dilute aqueous solution (15, 29), was left on the cells. The mutagenized cells were cultured until they reached confluence several days later. Plasmid DNA was then harvested from the cells by Hirt extraction (see Materials and Methods), and a portion of the DNA from each plate was transformed into the *E. coli* recipient strain MC1061 F'150 Kan *recA*. The frequency of LacI$^-$ mutants at each dose was determined by dividing the number of blue (LacI$^-$) colonies obtained by the total colony number (Fig. 2). Since 4,000 to 7,000 colonies were examined at each dose point in the curve, a detectable mutation frequency only appeared at doses of 40 μl/ml and higher. The mutation frequency rose steadily with dose. Killing effects of the mutagen were apparent in that a longer time was required for the heavily mutagenized cells to reach confluence.

A dose of 100 μg per ml of NMU (1 nM) was chosen for mutation collection since this dose gives a high mutation frequency while permitting >50% cell survival. Plates of 1.11 cells were mutagenized as in the dose-response experiment, and plasmid DNA was harvested when the plates became confluent 3 to 5 days later. Plasmid DNA was transformed into MC1061 F'150 Kan *recA* cells in which LacI$^-$ colonies were scored. In this way a collection of 270 NMU-induced mutations was derived from 78 independent plates of human cells, which yielded a total of 131,411 colonies. Thus, a LacI$^-$ mutant frequency of 2.1×10^{-3} was obtained. This induced mutation frequency is more than 300-fold above the spontaneous background of 6.4×10^{-6} for 1.11. Plasmid DNA from each of the 270 LacI$^-$ mutant

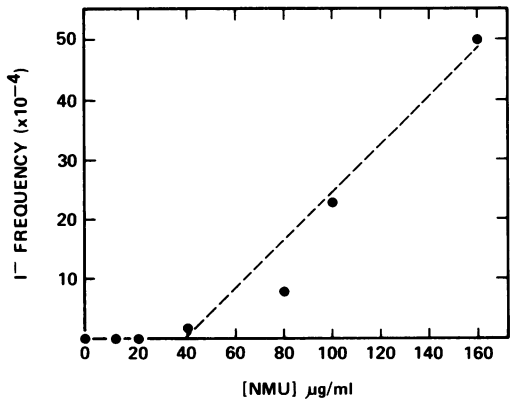

FIG. 2. NMU dose-response curve. Plasmid DNA from several plates of 1.11 cells treated with NMU was used to determine each dose point. See text.

TABLE 3. Induction of mutation by NMU[a]

Plasmid	Treatment	No. of colonies $\times 10^5$	No. LacI$^-$	LacI$^-$ frequency	Induction
pMCil DNA in DMEM	None	6.5	15	2.3×10^{-5}	2.9×
	100-μg/ml NMU	5.5	37	6.7×10^{-5}	
pMCil in MC1061 F'150 Kan *recA*	None	4.6	13	2.8×10^{-5}	2.6×
	100-μg/ml NMU	4.3	32	7.4×10^{-5}	
pMCil in 293 human cells (line 1.11)	None	1.5	1	6.4×10^{-6}	305×
	100-μg/ml NMU	1.3	270	2.1×10^{-3}	

[a] Upper: pMCil DNA was treated with NMU in vitro and transformed into MC1061 F'150 Kan *recA*. The number of bacterial colonies scored with and without NMU treatment and the number and frequency of LacI$^-$ colonies are given. DMEM, Dulbecco modified eagle medium. Middle: MC1061 F' Kan *recA* cells carrying pMCil were mutagenized with NMU and grown up. Plasmid DNA was purified from them and transformed into MC1061 F'150 Kan *recA*. Number of colonies and number and frequency of LacI$^-$ mutants were scored. Lower: Human 1.11 cells carrying pMCil were treated with 100 μg of NMU per ml and allowed to become confluent. Plasmid DNA was purified and returned to *E. coli*, in which colony number and number and frequency of LacI$^-$ mutants were scored. The level of induction of mutation by each treatment is shown.

bacterial colonies was prepared, digested with *Eco*RI, and run on an agarose gel. This analysis revealed that all the *lacI* plasmids appeared to be identical in size to wild-type pMCil DNA. In particular, the 1.7-kb *Eco*RI fragment carrying the *lacI* gene was wild type in length for all 270 *lacI* mutations. Therefore, all the NMU-induced mutations appeared to be of the point mutation class.

The following experiment was done to provide evidence that the mutations were formed and completed in the human cell and did not result from the processing of damaged DNA in the *E. coli* recipient. pMCil DNA was treated with 100 μg of NMU per ml in tissue culture medium (Dulbecco modified Eagle medium) for 2 h at 37°C. The mutagenized DNA was then ethanol precipitated and transformed into *E. coli* MC1061 F'150 Kan *recA*, and LacI$^-$ colonies were scored. The direct exposure of naked pMCil DNA to the full mutagenic dose of NMU surely overestimates the amount of damage that would be present in pMCil DNA mutagenized while inside the human cells, complexed with chromatin, and left for several days, during which multiple rounds of replication would be expected to have occurred. Yet, pMCil DNA directly treated with NMU led to less than a threefold increase in mutation frequency (Table 3). This value should be compared with the 300-fold increase after treatment of human cells with the same dose. The low mutagenicity of alkylated DNA transformed directly into *E. coli* argues against any major role for the bacterial cell in creating mutations in response to preexisting damage in the incoming plasmid DNA. Rather, the mutations appear to be formed and completed in the human cell.

The response of *E. coli* cells directly mutagenized with NMU was also monitored. An exponentially growing culture of approximately 10^7 MC1061 F'150 Kan *recA* cells carrying wild-type pMCil was exposed to 100 μg of NMU per ml in 5 ml of LB growth medium. After 2 h the culture was diluted to 15 ml and grown to saturation overnight. Plasmid DNA was extracted from the cells and used to transform MC1061 F'150 Kan *recA*. Blue (LacI$^-$) colonies were scored. The mutation frequency of the plasmid DNA mutagenized while in *E. coli* was elevated less than threefold above background (Table 3). This result indicates that the bacterial cells are relatively insensitive (approximately 100 times less sensitive) to an NMU dose that produced a very high mutation frequency in human cells and lends further weight to the argument that the NMU mutations we observed in the EBV shuttle system were completed in the human cells.

Analysis of mutations. Some 225 NMU-induced *lacI* mutations were subjected to genetic analysis. The mutations were first crossed by genetic recombination to the F' *lac pro* episome of GM1 (9). Nonsense mutations were detected by screening against a set of nonsense suppressors. We found 34 independent nonsense mutations among the 225 *lacI* mutations (15%). These were characterized and assigned to specific loci as previously described (4). Table 4 and Figure 3 portray the results. Of the 34 nonsense mutations, 33 resulted from G · C to A · T transitions. To test whether A · T to G · C transitions, which do not generate nonsense codons, were also induced by NMU, we examined a set of nonsuppressible mutations, using only one randomly chosen mutation from each culture to assure independence. This set of 56 mutations could potentially contain all categories of base substitutions, as well as frameshifts and larger deletions and insertions.

Table 4 displays the distribution of 27 of the nonsuppressible mutations induced by NMU, as revealed by genetic mapping and DNA sequence determination. Several distinct clusters occur, including one prominent hotspot. Table 4 also shows the sequence changes found in these 27 mutants. Of 27 mutations, 26 represent G · C to A · T transitions, the single exception being an A · T to G · C transition. The most prominent cluster represents G · C to A · T changes at a GGG sequence. Seven of the mutations in this cluster occur at the central guanine, and three occur at an adjacent guanine. Some of the other clusters also occur at monotonous runs of Gs. Among the sequenced missense mutations, a significant preference was shown for the mutational site to have a guanine residue 5′ to the guanine affected by the mutation (Table 5).

DISCUSSION

EBV shuttle. We described a system in which a bacterial gene, *lacI*, residing on an autonomously replicating EBV-based vector in human cells can be used to monitor mutagenesis. The vector can be easily recovered from the human cells after mutagenesis, and the mutations can be rapidly detected in bacteria. Using bacterial genetics in combination with biochemical techniques, sizable collections of mutations can be characterized at the DNA sequence level. A key feature of the shuttle system described here is the low spontaneous mutation frequency of the target gene in human cells. This frequency is less than 10^{-5} per gene or less than 10^{-8} per base pair, putting it in the range of true spontaneous mutation for both human (20) and bacterial (23) genes.

TABLE 4. DNA sequences at sites of NMU-induced mutations formed in human cells

Site	Amino acid position	No. of mutations	Sequence[a]	Base pair substitutions[b] G·C to A·T	A·T to G·C	G·C to T·A
Nonsense						
A5	18	1	GTCTGATAA	*		
O11	78	1	ATTTGCGAC	*		
O13	89	1	AGTTGATCG	*		
O17	117	1	CGTTGCGCG	*		
A15	131	1	TCCTGGTCA	*		
A16	153	2	GGCTGGTCA	*		
O21	181	2	ATTTGCTGG	*		
A21	201	1	GGCTGGCAT	*		
U5	201	3	GCAGGCATA	*		
O24	209	2	ATTTGATTG	*		
A23	211	1	GGCTGAATT	*		
U6	220	2	ACTGGAGTG	*		
O27	227	3	TGTTGAAAA	*		
O28	228	2	GTTTGTTGA	*		
O29	231	1	ATCTGCATG	*		
A26	248	4	ATCTGAGCG	*		
A28	269	1	TAGAGCCAT			*
A31	291	2	TCCTGTTTG	*		
A33	309	1	CCCTGAGAG	*		
A35	317	2	AGCTGATTG	*		
Missense						
	10	1	TGTCGCAGA	*		
	19	1	GTCTGGCAA	*		
	38	2	AAAAGTGGA	*		
	49	1	GTAAGGGTT	*		
	49	2	TAAGGGTTG	*		
	53	2	CGTGGCACA	*		
	56	1	GTTGACCGC		*	
	57	1	ACTGGCGGG	*		
	58	1	GCGGGCAAA	*		
	272	7	AGTGGGATA	*		
	272	3	GTGGGATAC	*		
	288	1	TGGTGGTAG	*		
	297	3	GCTGGGGCA	*		
	297	1	CTGGGGCAA	*		

[a] The base that changes is underlined. The type of base pair substitution is indicated at the right.

[b] There were no G · C to C · G, A · T to T · A, or A · T to C · G substitutions.

The low spontaneous mutation frequency observed in the EBV shuttle system contrasts with the higher frequencies characteristic of SV40-based shuttle systems. Spontaneous mutation frequencies for those systems typically range from 10^{-2} to 4×10^{-4} per gene or 10^{-5} to 4×10^{-7} per base pair, depending on the vector and cell type used (1, 3, 16, 27, 30, 34). The mutation frequency for the 1,000-bp *lacI* gene in the described clonal EBV system is 6×10^{-9} per base pair, a gain of more than 50-fold over the best SV40-based shuttle system available (18).

Previous studies with SV40-based shuttle systems concluded that the high frequency of spontaneous mutations recovered after passage of the vectors through mammalian cells was primarily due to damage incurred in the DNA while in the mammalian cell (16, 26). Possibly, newly transfected DNA is a particularly accessible substrate for chemical and enzymatic attack (24). The results reported here with an EBV shuttle system support this interpretation. When transfected cells are harvested as a population derived from many individual clones, a high mutation frequency similar to

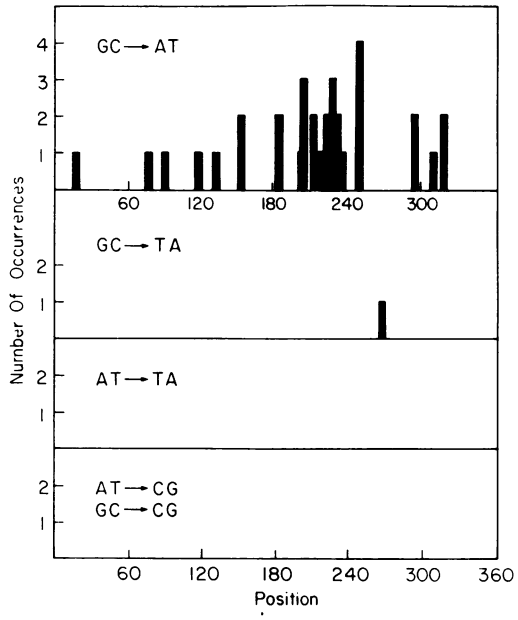

FIG. 3. Mutational specificity of NMU-induced nonsense base substitutions in human cells. The mutational change is shown in the upper left corner of each box. The position of the mutation along the 360-amino acid sequence of the *lac* repressor (coded for by *lacI*) is shown on the horizontal axis. The height of each bar represents the number of mutations at that site. The *lacI* nonsense system monitors 83 different single-base substitutions at UAG, UAA, and UGA sites (30 transitions and 53 transversions). Only the A · T to G · C change is not detected.

that of transiently replicating systems was observed. This result is interpretable as a mixture of cells whose plasmid populations incurred no mutations during transfection and cells which harbor plasmids which did mutate with transfection and perpetuate these mutant plasmids. These results are fortified by the findings with individual clones, in which approximately 90% of the clones apear to be free of mutation, while in the remainder recurrent mutations are repeatedly harvested at a high frequency. The high mutation frequency seen in all the populations and in a few isolated clones is independent of any SV40 replication, since T antigen was not present. When the same EBV vectors are placed in a transient replication situation using their SV40 origin of replication in the presence of T antigen, a high mutation frequency is again observed.

The above results argue that the high mutation frequency that is seen generally in transiently replicating plasmids and in pooled populations or sporadic isolated clones with stably replicating vectors such as those based on EBV is the result

TABLE 5. Effect of surrounding sequences[a]

5′ base	No. of mutations	Mutated base	3′ base	No. of mutations
G	16	G	G	6
A	2	G	A	4
T	0	G	T	4
C	1	G	C	5

[a] Mutational sites are tabulated with respect to the 5′ and 3′ bases flanking the site for G to A changes. Nineteen random missense mutations are analyzed. The seven occurrences at the hotspot, although preceded by a 5′ G, are excluded from the analysis to avoid possible bias.

of transfection, not replication. Unlike vectors based on EBV, vectors based on bovine papilloma virus show high mutation frequencies, even in clonal lines, as has been observed by us and others (see, for example, reference 2). The mutations, mainly rearrangements, appear to be due to the inherent instability of bovine papilloma virus vectors. Drinkwater and Klinedinst (7) examined populations of lymphoblastoid cells carrying EBV-based shuttle vectors for spontaneous mutation frequency. In this case, the frequency was approximately 6.6×10^{-5} per gene, and the mutations were mainly attributed to transfection. This frequency is fivefold below that observed for pooled populations of 293 cells carrying EBV-*lacI* vectors. These authors also reported data consistent with an increase in mutation frequency over time (7), in contrast to our own results with 1.11. The reason for these differences may be related to differences in the vectors, method of DNA introduction, or human cell lines.

The absence of an elevated mutation frequency for EBV-*lacI*-carrying clones such as the 1.11 cell line described here makes them ideal for studies of induced mutation. The ease of mutation collection and analysis is in striking contrast to approaches confronting chromosomal genes directly. However, it is necessary to verify that the data obtained with the shuttle vector approach will be applicable to true chromosomal genes. Although the EBV vectors are presumably complexed into chromatin and are completely dependent on host enzymes for replication and repair, it is conceivable that they will differ in some ways from authentic chromosomes with respect to mutagenesis. The data currently available regarding the type of mutation made by a given mutagen in genes embedded in chromosomes are in agreement with data derived with shuttle vectors. For example, the sequence specificities of eight UV-induced mutations rescued from the chromosome by excision (10) are in agreement with the sequence changes derived for 53 UV-induced mutations analyzed with an autonomously replicating SV40-based shuttle system (18). Susceptibility to mutation of a gene carried on an EBV shuttle vector compared with that of a chromosomal gene also seems to be similar. Drinkwater and Klinedinst (7) found that the mutation frequency induced by ethyl nitrosourea for the gene they monitored on an EBV shuttle vector paralleled that of the chromosomal *hprt* locus.

The low spontaneous background of the EBV-*lacI* shuttle system gives it great sensitivity for the examination of the molecular biology of mutation in human cells at doses that are biologically relevant. The negligible response of bacteria to a dose of NMU that produced dramatic results in human cells points up the relevance of using human cells in mutagen testing and screening. Convenient tester systems are being built with the EBV-*lacI* shuttle in a variety of normal and mutant human cells. The sensitivity of these systems may be improved by alterations which raise the number of vector molecules that can be conveniently screened for mutation. Some of these alterations involve raising the transformation efficiency of the bacterial recipient (12) and the copy number per cell of the vector (8).

NMU-induced mutations. The power of the EBV shuttle system for detection and analysis of mutation in human cells is apparent in the NMU study. Mutation was readily induced more than 300-fold above background by mutagenizing living cells. Therefore, analysis of the mutations produces the mutagenic specificity of NMU with little interference from the spontaneous background. Analysis of 225 mutations was readily accomplished by exploiting the *lacI* genetic system. Mutations were mapped to within 5 to 10 bp by genetic techniques. The *lacI* nonsense system was used to assign the sequence changes of the 34 nonsense mutations in the collection. The strength of the nonsense system is its ability to rapidly assign sequence changes for five of six possible base changes at 83 sites throughout the gene. To monitor the sixth change and to examine more closely features such as the hotspot uncovered by genetic mapping, we used direct DNA sequencing. The sequencing, guided by precise genetic localization of the mutations, added another 27 DNA sequence changes to the study.

A total of 33 of 34 nonsense mutations and 26 of 27 missense mutations involve the G · C to A · T transition. All the mutations involve single base pair substitutions. The overwhelming preference for NMU is, therefore, the G · C to A · T transition (97%). NMU has been well documented as an alkylating agent which can add a methyl group at many positions on DNA (22, 35). The mutagenic lesion for the G · C to A · T transition is presumably addition of the alkyl group to the O^6 position of guanine, leading to mispairing with thymine at the first round of replication and fixation of the full transition at the next round (reviewed in reference 6). Numerous other lesions are made in DNA by alkylating agents, notably alkylation of the O^4 position of thymine. This lesion could also potentially cause mispairing, leading to A · T to G · C transitions (36). One example of this change, which is the only sequence change not detected by the nonsense system, was found among 28 missense mutations sequenced. Therefore, it is possible that this change plays a minor role in NMU mutagenesis in vivo.

Examination of the DNA sequences surrounding the missense mutational sites reveals a strong preference for guanine at the 5′ side of the mutated base. In this regard, it is of interest that the one apparent hotspot for NMU mutation within *lacI* occurs within a run of three consecutive guainine residues. Of 10 mutations, 7 at the site are at the central G, while 3 others are at the 3′ G. Possibly, guanine residues are more reactive with NMU when they occur adjacent to other purines, especially guanine residues. It is known that guanine is the most chemically reactive nucleophilic center in DNA (35).

This study has also provided support for the mutational theory of cancer in connection with the experiments of Barbacid and colleagues (38, 43). In those studies, NMU was used to induce mammary carcinomas with great efficiency in Buf/N rats. Tumor tissue from most of the animals contained activated Ha-*ras* genes. In 48 of 48 cases, the altered *ras* genes contained a G · C to A · T transition in the second base of the codon for amino acid 12. The mutations were interpeted to be direct products of interaction of the relevant guanine bases with NMU. This interpretation was based on the short half-life of NMU and the likelihood that it would produce transitions in mammalian cells based on results with alkylating agents in bacteria. The present study provides unequivocal evidence that NMU induces the G · C to A · T transition efficiently and as its dominant mode of mutagenesis in mammalian cells. The results thus support the conclusions of Barbacid and colleagues and the mutational theory of cancer.

It is of interest that the G · C to A · T transition observed in codon 12 occurred in all cases at the second guanine of the GGA codon, even though a change at the first guanine would also have produced an activated *ras* gene (33, 39). This specificity may be affected by the preference for NMU-induced mutation at bases preceded by a guanine, a juxtaposition not present for the initial guanine of codon 12 (GCT GGA). In contrast, codon 61, which is often mutated in activated *ras* alleles, is not affected by NMU. As noted by

Zarbl et al. (43), this codon (CAA) could give a nonsense but not a missense mutation via a G · C to A · T transition. However, amino acid 61 is mutated at an A · T base pair when mammary carcinomas are induced in Buf/N rats by dimethyl benzanthracene, a mutagen entirely distinct from NMU (43). Also, an A · T to T · A change at codon 61 was seen in over 90% of skin tumors induced in mice by dimethyl benzanthracene (25). Further evidence favoring the mutational theory of cancer would result if the mutational specificity of DMBA in mammalian cells, currently unknown, was consistent with a missense change at codon 61 rather than codon 12. Determination of the mutational specificity of dimethyl benzanthracene in human cells with the EBV-*lacI* shuttle is currently in progress. The availability of the EBV-*lacI* shuttle system should permit the mutagenic specificity of any mutagen to be conveniently determined in human cells.

ACKNOWLEDGMENTS

We are grateful to Bill Sudgen for providing p220.2 and generous advice, Robert Tjian for pRSV-1609 plasmid DNA, and Jo Ann Katheiser for secretarial assistance.

R.B.D. was supported by training grant GM07599 from the National Institutes of Health. This work was supported by a grant from the National Institutes of Health to J.H.M. and grants from the National Cancer Institute (CA33056) and Chicago Community Trust/Searle Scholar's program to M.P.C.

LITERATURE CITED

1. **Ashman, C. R., and R. L. Davidson.** 1984. High spontaneous mutation frequency in shuttle vector sequences recovered from mammalian cellular DNA. Mol. Cell. Biol. **4**:2266–2272.
2. **Ashman, C. R., and R. L. Davidson.** 1985. High spontaneous mutation frequency of BPV shuttle vector. Somatic Cell Mol. Genet. **11**:499–504.
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. USA **80**:3015–3019.
4. **Coulondre, C., and J. H. Miller.** 1977. Genetic studies of the *lac* repressor. III. Additional correlation of mutational sites with specific amino acid residues. J. Mol. Biol. **117**:525–567.
5. **Coulondre, C., and J. H. Miller.** 1977. Genetic studies of the *lac* repressor. IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli*. J. Mol. Biol. **117**:577–606.
6. **Drake, J. W., and R. H. Baltz.** 1976. The biochemistry of mutagenesis. Annu. Rev. Biochem. **45**:11–37.
7. **Drinkwater, N. R., and D. K. Klinedinst.** 1986. Chemically induced mutagenesis in a shuttle vector with a low-background mutant frequency. Proc. Natl. Acad. Sci. USA **83**:3402–3406.
8. **DuBridge, R. B., M. Lusky, M. R. Botchan, and M. P. Calos.** 1985. Amplification of a bovine papilloma virus-simian virus 40 chimaera. J. Virol. **56**:625–627.
9. **Farabaugh, P. J., U. Schmeissner, M. Hofer, and J. H. Miller.** 1978. Genetics studies of the *lac* repressor. VII. On the molecular nature of spontaneous hotspots in the *lacI* gene of *E. coli*. J. Mol. Biol. **126**:847–857.
10. **Glazer, P. M., S. N. Sarkar, and W. C. Summers.** 1986. Detection and analysis of UV-induced mutations in mammalian cell DNA using a λ phage shuttle vector. Proc. Natl. Acad. Sci. USA **83**:1041–1044.
11. **Graham, F. L., J. Smiley, W. C. Russell, and R. Nairn.** 1977. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. J. Gen. Virol. **36**:59–72.
12. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166**:557–580.
13. **Hauser, J., M. M. Seidman, K. Sidur, and K. Dixon.** 1986. Sequence specificity of point mutations induced during passage of a UV-irradiated shuttle vector plasmid in monkey cells. Mol. Cell. Biol. **6**:277–285.
14. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cultures. J. Mol. Biol. **26**:365–369.
15. **Jensen, E. M., L. J. La Polla, P. E. Kirby, and S. R. Haworth.** 1977. *In vitro* studies of chemical mutagens and carcinogens. I. Stability studies in cell culture medium. J. Natl. Cancer Inst. **59**:941–944.
16. **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.
17. **Lebkowski, J. S., S. Clancy, and M. P. Calos.** 1985. SV40 replication in adenovirus-transformed human cells antagonizes gene expression. Nature (London) **317**:169–171.
18. **Lebkowski, J. S., S. Clancy, J. H. Miller, and M. P. Calos.** 1985. The *lacI* shuttle: rapid analysis of the mutagenic specificity of ultraviolet light in human cells. Proc. Natl. Acad. Sci. USA **82**:8606–8610.
19. **Lebkowski, J. S., J. H. Miller, and M. P. Calos.** 1986. Determination of DNA sequence changes induced by ethyl methanesulfonate in human cells, using a shuttle vector system. Mol. Cell. Biol. **6**:1838–1842.
20. **Lewin, B.** 1980. Gene expression: eukaryotic chromosomes, vol. 2, 2nd ed., 142–188. John Wiley & Sons, Inc., New York.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **Margison, G. P., and P. J. O'Connor.** 1979. Nucleic acid modification by N-nitroso compounds, p. 111–159. *In* P. L. Grover (ed.), Chemical carcinogens and DNA, vol. 1. CRC Press, Inc., Boca Raton, Fla.
23. **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.
24. **Miller, J. H., J. S. Lebkowski, K. S. Greisen, and M. P. Calos.** 1984. Specificity of mutations induced in transfected DNA by mammalian cells. EMBO J. **13**:3117–3121.
25. **Quintanilla, M., K. Brown, M. Ramsden, and A. Balmain.** 1986. Carcinogen-specific mutation and amplification of Ha-*ras* during mouse skin carcinogenesis. Nature (London) **322**:78–80.
26. **Razzaque, A., S. Chakrabarti, S. Joffee, and M. Seidman.** 1984. Mutagenesis of a shuttle vector plasmid in mammalian cells. Mol. Cell. Biol. **4**:435–441.
27. **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. USA **80**:3010–3014.
28. **Rio, D., S. G. Clark, and R. Tjian.** 1985. A mammalian host-vector system that regulates expression and amplification of transfected genes by temperature induction. Science **277**:23–28.
29. **Sanders, F. K., and B. O. Burford.** 1967. Morphological conversion of cells *in vitro* by N-nitrosomethylurea. Nature (London) **213**:1171–1173.
30. **Sarkar, S., U. B. Das Gupta, and W. C. Summers.** 1984. Error-prone mutagenesis detected in mammalian cells by a shuttle vector containing the *supF* gene of *Escherichia coli*. Mol. Cell. Biol. **4**:2227–2230.
31. **Schmeissner, U., D. Ganem, and J. H. Miller.** 1977. Genetic studies of the *lac* repressor. II. Fine structure deletion map of the *lacI* gene, and its correlation with the physical map. J. Mol. Biol. **109**:303–326.
32. **Schmeissner, U., D. Ganem, and J. H. Miller.** 1977. Revised gene-protein map for the *lacI* gene-*lac* repressor system. J. Mol. Biol. **117**:572–575.
33. **Seeburg, P. H., W. W. Colby, J. S. Hayflick, D. J. Capon, D. V. Goeddel, and A. D. Levinson.** 1984. Biological properties of human c-Ha-*ras*1 genes mutated at codon 12. Nature (London) **312**:71–75.
34. **Seidman, M. M., K. Dixon, A. Razzaque, R. J. Zagursky, and M. L. Berman.** 1985. A shuttle vector plasmid for studying carcinogen-induced point mutations in mammalian cells. Gene **38**:233–237.
35. **Singer, B., and D. Grunberger.** 1983. Molecular biology of

mutagenes and carcinogens. Plenum Publishing Corp., New York.
36. **Singer, B., S. J. Spangler, H. Fraenkel-Conrat, and J. T. Kusmierek.** 1986. O^4-methyl, -ethyl, or -isopropyl substituents on thymidine in poly (dA-dT) all lead to transitions upon replication. Proc. Natl. Acad. Sci. USAp **83**:28–32.
37. **Sugden, B., K. Marsh, and J. Yates.** 1985. A vector that replicates as a plasmid and can be efficiently selected in B-lymphoblasts transformed by Epstein-Barr virus. Mol. Cell. Biol. **5**:410–413.
38. **Sukumar, S., V. Notario, D. Martin-Zanca, and M. Barbacid.** 1983. Induction of mammary carcinomas in rats by nitroso-methylurea involves activation of H-*ras*-1 locus by single point mutations. Nature (London) **306**:658–661.
39. **Varmus, H. E.** 1984. The molecular genetics of cellular oncogenes. Annu. Rev. Genet. **18**:553–612.
40. **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.
41. **Yates, J., N. Warren, D. Reisman, and B. Sugden.** 1984. A *cis*-acting element from the Epstein-Barr viral genome that permits stable replication of recombinant plasmids in latently infected cells. Proc. Natl. Acad. Sci. USA **81**:3806–3810.
42. **Yates, J., N. Warren, and B. Sugden.** 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature (London) **313**:812–815.
43. **Zarbl, H., S. Sukumar, A. V. Arthur, D. Martin-Zanca, and M. Barbacid.** 1985. Direct mutagenesis of Ha-*ras*-1 oncogenes by *N*-nitroso-*N*-methylurea during initiation of mammary carcinogenesis in rats. Nature (London) **315**:382–385.