

## Radon-induced Deletions in Human Cells: Role of Nonhomologous Strand Rejoining<sup>1</sup>

Louise H. Lutze,<sup>2</sup> Richard A. Winegar, Rick Jostes, Fredrick T. Cross, and James E. Cleaver

Laboratory of Radiobiology and Environmental Health, University of California, San Francisco, California 94143-0750 [L. H. L., J. E. C.]; SRI International, Menlo Park, California 94025 [R. A. W.]; and Pacific Northwest Laboratory, Richland, Washington 99352 [R. J., F. T. C.]

### Abstract

Radon is a ubiquitous inhaled human carcinogen that is thought to be the largest single natural source of human exposure to radiation. We report that a freely replicating episome in human cells exposed to radon gas underwent mutagenic changes, a high proportion of which were large deletions involving many thousands of base pairs. These deletions were not randomly distributed but started and ended in defined regions as if caused by the passage of a single  $\alpha$ -particle track through a coiled chromatin structure. The sizes appeared to be defined by structural features of chromatin: the minimum size was 2435 base pairs, and the maximum size was 8051 base pairs, close to the upper limit that would leave intact the plasmid sequences required for selection in bacteria. Ends were rejoined by nonhomologous recombination involving up to 6 base pairs of homology. This process may not be confined to the repair of exogenously induced double-strand breaks but may be used for rejoining free DNA ends generated by a variety of cellular processes. The mechanism of  $\alpha$ -particle deletion mutagenesis may account for the high relative biological effectiveness of radon irradiation for many end points and its consequences for lung carcinogenesis.

### Introduction

Radon has recently been recognized as a major source of human exposure to background radiations. It is estimated that 5,000 to 25,000 lung cancer deaths per year in the United States, including one-half of the lung cancers in nonsmokers, can be attributed to radon exposure (1). Radon is a gaseous  $\alpha$ -particle emitter produced during the radioactive decay of uranium, which is released from soils, rocks, water, and building materials and is contained in natural gas and other fossil fuels. Diffusion of the gas results in detectable levels of radon in homes and can require costly mitigation procedures (1). The route of human exposure is through inhalation, which results in the deposition of solid radioactive daughter elements in the lungs, with consequent high exposure of lung epithelial cells to densely ionizing tracks from  $\alpha$ -particles. Passage of  $\alpha$ -particles through cells can give rise to mutations and malignant transformation (1). It is important, therefore, to understand the mechanisms of interaction of radon with human cells in as much molecular detail as possible. We have designed a shuttle vector system to examine the molecular consequences of radon exposure to human cells, with particular emphasis on the molecular mechanisms by which  $\alpha$ -particle-induced deletions are produced and DNA is rejoined. A previous study by others using the vector pZ189 with a 200-base pair target showed that

point mutations induced by radon are predominantly GC:AT transitions (2). This vector is not suitable for the study of deletion mutagenesis, however, because the target gene is too small and the vector is subject to rearrangements when introduced into the cell for each experiment. Our vector, pHAZE, has the entire 3.1-kilobase *Escherichia coli*  $\beta$ -galactosidase gene (*lacZ*) as a target for mutagenesis, an EBV oriP,<sup>3</sup> which allows maintenance of the vector as an episome in suitable mammalian cells, and sequences from pBR327 to allow replication and selection in a bacterial host for mutant recovery and identification. We made a clonal derivative of a human diploid lymphoblastoid cell line, Raji F1, which maintains the vector at  $\sim 80$  copies/cell with a very low spontaneous mutation frequency ( $2.9 \times 10^{-5}$ ). Initial studies in which X-rays were used as the mutagen indicated that this system can detect both point mutations and deletions, with deletion sizes extending up to 8.3 kilobases (3). From radon exposure we have determined (a) the proportion of deletions among all mutants, (b) the deletion size distribution, (c) the presence of preferred regions in the plasmid for deletion end points, and (d) the DNA sequences at deletion junctions.

### Materials and Methods

F1 cells were isolated and maintained as described previously (3). They were exposed to radon gas and its progeny at Pacific Northwest Laboratory, Richland, WA (4). After a 24-h recovery period, cells were harvested and plasmid DNA was recovered by alkaline lysis and ion-exchange chromatography (5). The purified plasmid DNA was transformed into *E. coli* strain WM1100 ( $\Delta lacZ$ , *recA*<sup>-</sup>) by electroporation. Mutant plasmids were identified and isolated as described previously (3). RFLP mapping was used to determine the types of mutations induced. The junction regions of mutants containing large rearrangements were sequenced by means of the dideoxy chain termination method with a T7 sequencing kit (Pharmacia LKB, Piscataway, NJ) and primers specifically synthesized to anneal throughout the target region.

### Results and Discussion

RFLP analysis of mutants revealed that radon had induced point mutations (defined in this study as alterations that showed no change on RFLP analysis, meaning that <50 base pairs were affected), deletions, and insertions in pHAZE molecules. Most of the changes detected were deletions (64%), in contrast to only 13% deletions in our previous study with X-rays (3). Further mapping revealed that these deletions ranged in size from 2435 base pairs to 8051 base pairs (Table 1), close to the maximum detectable with this system (Fig. 1A). There was a continuum of deletion sizes within this range, but

Received 7/27/92; accepted 8/10/92.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

<sup>1</sup> This work was supported by Office of Health and Environmental Research, United States Department of Energy, Contracts DE-AC03-76-SF01012 and DE-AC06-76-RL01830 and by NIH Contract R29 GM46563-01 (R. A. W.).

<sup>2</sup> To whom requests for reprints should be addressed.

<sup>3</sup> The abbreviations used are: EBV, Epstein-Barr virus; oriP, origin of replication; RFLP, restriction fragment length polymorphism.

unexpectedly there were none in the range between 2.4 kilobases and those classified as point mutations, suggesting that deletion sizes may be influenced by the conformational and functional properties of the DNA region. No mutants containing complex inter- or intraplasmid rearrangements were recovered, although mutants containing intraplasmid rearrangements have been recovered in a parallel study of restriction enzyme-induced damage (6).

All deletions removed part of the *lacZ* region of the plasmid but did not extend into the bacterial sequences that are required for recovery of the plasmids in *E. coli*. All but two deletions extended into or encompassed the EBV oriP region, which is the mammalian origin of replication for the plasmid (Fig. 1A). Consequently, these mutants would not be maintained long-term in Raji cells, but their recovery means that we are able to detect deletions at early times after their formation without the complicating factors of replication and selection.

Deletion junctions appeared to be clustered (Fig. 1B); 82% of deletions started in the first 1.5 kilobases of the plasmid, and 71% ended between positions 5.5 kilobases and 7.2 kilobases of pHAZE. This 1.7-kilobase length of the plasmid contains the sequences for replication in mammalian cells and the transcriptionally active gene for hygromycin resistance. The clustering of deletion sites may reflect greater accessibility of these regions to radiation damage owing to the more open structure of chromatin at these sites. Alternatively, the folding of the plasmid in the cell might bring these two regions into close proximity, so that the track of one  $\alpha$ -particle could cause a break in both regions. The proximity of the ends would then favor their rejoining.

We found limited sequence homology at the junction regions in 65% of the mutants containing large rearrangements (Table 1; Fig. 2). This homology involved up to 6 base pairs and would not be predicted to occur by chance ( $P < 0.0001$ ,  $\chi^2$  analysis; Table 2). In two cases (12%), a single base pair (a C in the strand that was sequenced) was inserted at the junction. Only one mutant (R21) had a large insertion of DNA, ~6 kilobases in size, which was not derived from pHAZE and was presumably genomic in origin. This mutant also had a large deletion of pHAZE DNA (7.1 kilobases) at the insertion site. The remaining mutants (17%) had no homology in the sequences at the junction region.

Table 1 Spectrum of mutations induced in pHAZE by exposure to radon in vivo

Mutant	Size of deletion (base pairs)	Junction region	
		Insert (no. of bases) <sup>a</sup>	Terminal homology (no. of bases) <sup>a</sup>
R1	6245		0
R5	4559		5 (ACGGC)
R6	5511		3 (CAC)
R10	2435		4 (GGCG)
R13	5479	1 (C)	
R14	3760		1 (A)
R15	3671		0
R18	8051		1 (A)
R21	7134	~6 kilobases	
R22	6854		3 (AGC)
R28	7577		2 (CA)
R29	6505	1 (C)	
R30	4717		3 (TCC)
R31	6063		3 (CGA)
R32	7913		6 (GCCCCG)
R33	5349		1 (C)
R45	3284		0

<sup>a</sup> The bases inserted or contained in the sequences with terminal homology are shown in parentheses.

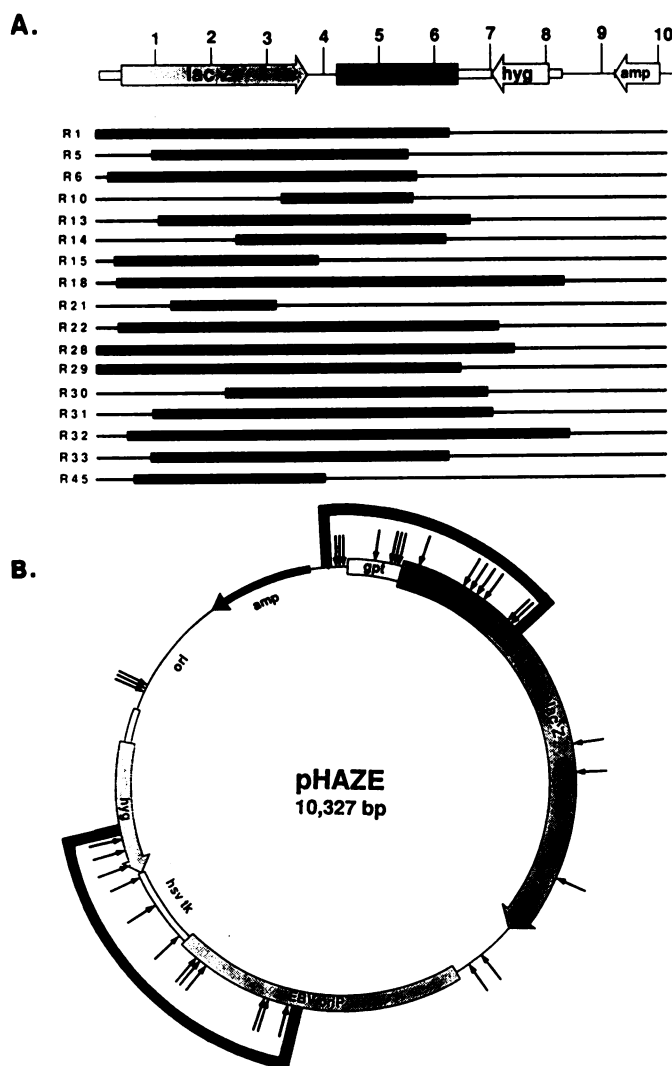


Fig. 1. A, radon-induced deletions in pHAZE. The vector, pHAZE, is shown linearized at the unique *Hind*III site at position 1; the numbers above the vector indicate its size in kilobases. The size and position of deletions induced by radon are shown for each mutant below. ■, regions of DNA deleted. *hyg*, hygromycin-resistance gene; *amp*, ampicillin-resistance gene. B, sites of junction regions in pHAZE. The sites of radon-induced double-strand breaks in pHAZE are indicated by arrows around the circumference of this map of the vector. The regions where junction regions are clustered are enclosed by thick lines. The bacterial sequences required for vector recovery in this system are shown as *ori* and *amp* in the diagram. *bp*, base pairs.

This study provides detailed information about the nature of deletions produced in human cells by exposure to an important environmental hazard. The large deletions produced by radon, their size range, the functions of the DNA regions deleted, and the limited homology of the junction sequences suggest a scenario for how deletions are produced by radon  $\alpha$ -particles and provide insight into general rules governing deletion mutagenesis.

The initial event resulting from radon exposure is likely to be a single  $\alpha$ -particle track through a pHAZE molecule with the production of more than one DNA double-strand break. Given the small target size of pHAZE and the radon dose used, it is extremely unlikely that pHAZE would be traversed by more than one  $\alpha$ -particle. It has been estimated<sup>4</sup> that in Raji cells

<sup>4</sup> T. E. Hui, personal communication.

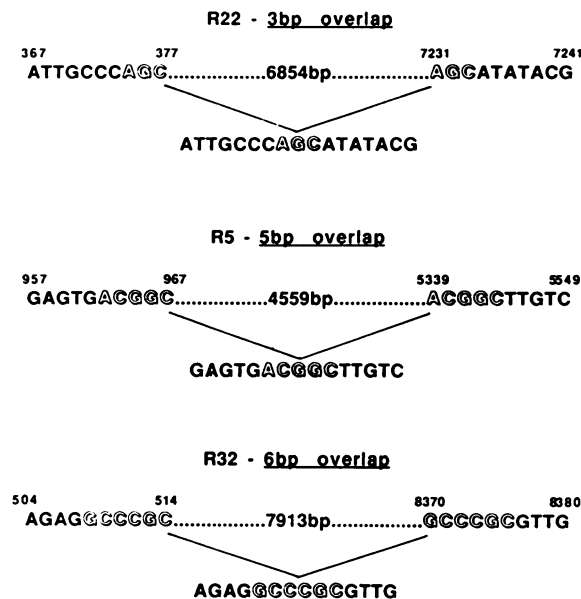


Fig. 2. Sequences at junction regions. The formation of deletion junctions with limited sequence homology is shown in three examples. The extent of the deletion between the junctions is given in base pairs (bp). The homologous sequences at the junction are shown as *hollow letters*. The positions of the junction regions in the parental plasmids are shown above the sequences.

the probability that a pHAZE molecule will be hit once by an  $\alpha$ -particle passing through the nucleus is  $1.6 \times 10^{-4}$ ; consequently, the probability that it will be hit twice is  $2.6 \times 10^{-8}$ . Furthermore, it has been shown that exposure of rodent cells to  $\alpha$ -radiation causes a linear increase in the number of chromosome aberrations per cell, in contrast to the dose-squared relationship observed after exposure to low-linear energy transfer radiation (7). This change in the shape of the dose-response curves is attributed to the production of the two breaks necessary for the induction of chromosomal exchanges by a single  $\alpha$ -particle track (7).

Despite the fact that one  $\alpha$ -particle appeared to be involved in the deletions, it is likely that two breaks within one DNA molecule are involved. If a single  $\alpha$ -particle track gave rise to only one double-strand break in pHAZE, then the production of the large deletions we detected would require extensive enzymatic degradation from the termini. This does not seem likely because most evidence suggests that termini in mammalian cells are very stable: 97% are rejoined within 15 nucleotides of the junction, and 83% are rejoined within 5 nucleotides of the termini (8). Parallel experiments with pHAZE damaged *in vivo* by restriction enzyme digestion indicated that DNA ends are rejoined between the restriction sites, with no evidence for extensive terminal degradation (6). Furthermore, the pattern of deletions we detected in pHAZE after  $\alpha$ -irradiation showed a continuum of sizes from 2.4 to 8 kilobases, with none smaller than 2.4 kilobases, which is more consistent with a two-hit mechanism. If the deletions had occurred by degradation from a single break, a continuous distribution of sizes, including some smaller than 2.4 kilobases, would be expected. In addition, the clustering of termini in two regions suggests that these regions were in close proximity and were traversed by one  $\alpha$ -particle track.

This interpretation could explain how the broken ends are brought together with limited sequence homology. The sequence that contains 6 base pairs of homology (GCCCCG)

occurs 5 times in pHAZE, the 5-base pair sequence (ACGGC) occurs 17 times, the 4-base pair sequence (GGCG) occurs 90 times, and each 3-base pair sequence found in junction regions occurs more than 200 times. The recombination events therefore are unlikely to be specifically directed to these sites, but when they do occur in proximity to the termini they may be used to facilitate rejoining. Since the pairing between sequences with very limited homology (especially of 1 and 2 nucleotides) would not be sufficient to hold duplexes together, it has been suggested that proteins may be involved in enhancing the stability of such end associations until they can be joined covalently (9, 10). The characteristics of the rejoined ends appear to fit the pattern described for nonhomologous recombination, which is an important cellular mechanism for rejoining many kinds of double-strand breaks.

It has been proposed that nonhomologous recombination in mammalian cells may occur at sites of inverted repeat sequences (11). None of the deletion junction regions examined in this study were at, or near, inverted repeat sequences in pHAZE. The oriP section of the vector contains a region of lengthy dyad symmetry and a noncontiguous region with 20 tandem copies of a 30-base pair inverted repeat sequence (12). All deletion junctions in oriP were at least 65 base pairs removed from these regions, and in most mutants these regions were completely deleted in pHAZE. While inverted repeat sequences may be implicated in the formation of spontaneous deletions in mammalian cells, they do not appear to play a role in the rejoining of breaks induced exogenously by  $\alpha$ -particles. We investigated the possibility that the rejoining might be mediated by either topoisomerase I or topoisomerase II. None of the seven putative mammalian topoisomerase I sites (13) were found to be involved in these junction regions, and there are no topoisomerase II sites in pHAZE. The EBV oriP region contained in pHAZE has recently been shown to be a matrix association region in Raji cells (14). This association of the plasmid with the nuclear matrix may influence both damage induction and repair in the vector, as suggested by the clustering of deletion end points in or near this region.

It has been shown that the severe combined immunodeficiency in mice (*scid*) is due to a defect in their ability to recombine immunoglobulin and T-cell receptor genes (15). These animals, and cell lines derived from them, also show an increased sensitivity to X-rays (16). This correlation suggests that

Table 2 Analysis of regions with homology<sup>a</sup>

Terminal homology (no. of bases)	P	Predicted no. of mutants	Observed no. of mutants
0	0.56	7.8	3
1	0.28	3.9	3
2	0.11	1.5	1
3	0.035	0.5	4
4	0.010	0.1	1
5	0.0032	0.045	1
6	0.00096	0.013	1

<sup>a</sup> The probability that a junction will contain homologous nucleotides was calculated by the equation

$$P(X) = (X + 1)(\frac{1}{4})^X(\frac{3}{4})^2$$

where  $X$  is the number of nucleotides of homology in the junction region,  $(X + 1)$  is the number of different ways in which chance identities could yield the specified homology,  $(\frac{1}{4})^X$  is the probability the  $X$  nucleotides match, and  $(\frac{3}{4})^2$  is the probability that nucleotides flanking the matching nucleotides do not match (19).  $\chi^2$  analysis of the data revealed that the number of mutants containing terminal homology differs significantly ( $P < 0.0001$ ) from that predicted for unbiased base composition at the junction regions.

a nonhomologous component of the recombinase system responsible for *V(D)J* recombination might also play a role in repair of exogenously induced double-strand breaks. Although the rejoining of the signal sequences recognized by the recombinase is very precise (9), the joining of the gene segments during immunoglobulin rearrangement occurs within a few nucleotides of the break; furthermore, there are zero to a few nucleotides of homology at the junction, and often some nucleotides, called N-regions, are inserted (9). This pattern is similar to that found in the junction regions of pHAZE and strengthens the possible correlation between the recombinase involved in immunoglobulin gene rearrangement and the repair of exogenously induced double-strand breaks in B-cells.

The types of junction regions found in this study have been found in other illegitimate recombination events. Spontaneous deletions detected at the *aprt* locus in cells are found to have 0–7 nucleotides of homology at the junction, with very little homology outside the two donor sequences (17). Gene rearrangements in diseases such as  $\alpha$ - and  $\beta$ -thalassemia, familial hypercholesterolemia, and Duchenne muscular dystrophy are most frequently deletions that also appear to be rejoined by this mechanism (11). The translocation of chromosome segments has been identified in several types of cancer; it is thought that translocation activates an oncogene and leads to the formation of a tumor (18). Translocation junctions show the same pattern of rejoining: limited homology at the junction and very little homology outside the junction; filler DNA is often present (19).

The most logical explanation for the rejoining of the  $\alpha$ -particle-induced deletions is therefore that they were repaired by a ubiquitous nonhomologous recombination mechanism. This mechanism, which has been demonstrated in monkey cells and extracts from frog eggs (8, 10, 20), appears to rejoin any pair of ends in a manner that is dependent on the structural features of the ends (8, 20). On the basis of these previous studies and our current results, we propose the following scheme for the rejoining of double-strand breaks produced by radiation: (a) the complex radiolytic products initially produced at the termini of radiation-induced DNA breaks are restored to intact nucleotides by terminal nuclease action; (b) blunt ends produced could then be substrates for terminal deoxynucleotide transferase. This enzyme catalyzes the template-independent addition of nucleotides to 3' ends in cells of lymphoid lineage, with a preference for the addition of G nucleotides (21). This would be consistent with our finding that additional nucleotides consisted of G-C base pairs; (c) ends that can be abutted (a 5' extension opposite a 3' extension or either extension opposite a blunt end) can be joined by a combination of fill-in synthesis and ligation. It has been shown that 50–70% of such ends in linearized DNA are rejoined by such a process (8, 20). This process would produce junctions that had no insertion of bases and appeared to arise by direct ligation of the broken ends, such as mutants *R1*, *R15*, and *R45* in the present study; (d) ends that can overlap, *i.e.*, two 5' extensions or two 3' extensions, can be joined by pairing between complementary nucleotides in the single-strand extensions. Short sequence pairing would explain the frequent appearance of one to six nucleotides of homology at the junctions found in this and other studies in mammalian cells (8, 11, 19, 20).

## Acknowledgments

We thank S. Wolff, J. P. Murnane, and D. Deen for their helpful comments on the manuscript. We also thank M. McKenney for editing the manuscript and P. Winegar for typing it.

## References

1. Evaluation of Occupational and Environmental Exposures to Radon and Radon Daughters in the United States, NCRP Report 78. Washington, DC: National Council on Radiation Protection and Measurements, 1984.
2. Jaberabansari, A., Dunn, W. C., Preston, R. J., Mitra, S., and Waters, L. C. Mutations induced in ionizing radiation in a plasmid replicated in human cells. II. Sequence analysis of  $\alpha$ -particle-induced point mutations. *Radiat. Res.*, 127: 202–210, 1991.
3. Lutze, L. H., and Winegar, R. A. pHAZE: a shuttle vector system for the detection and analysis of ionizing radiation-induced mutations. *Mutat. Res.*, 245: 305–310, 1990.
4. Jostes, R. F., Hui, T. E., James, A. C., Cross, F. T., Schwartz, J. L., Rotmensch, J., Atcher, R. W., Evans, H. H., Mencl, J., and Bakale, G. *In vitro* exposure of mammalian cells to radon: dosimetric considerations. *Radiat. Res.*, 127: 211–219, 1991.
5. Lutze, L. H., and Winegar, R. A. A quick and efficient method for the recovery of plasmid or viral DNA from mammalian cells. *Nucleic Acids Res.*, 18: 6150, 1990.
6. Winegar, R. A., Lutze, L. H., Rufer, J. T., and Morgan, W. F. Spectrum of mutations produced by specific types of restriction enzyme-induced double-strand breaks. *Mutagenesis*, in press, 1992.
7. Evans, H. H. Cellular and molecular effects of radon and other alpha particle emitters. In: G. Obe (ed.), *Advances in Mutagenesis Research*, pp. 28–52. New York: Springer-Verlag, 1991.
8. Roth, D. B., and Wilson, J. H. Nonhomologous recombination in mammalian cells: role for short sequence homologies in the joining reaction. *Mol. Cell. Biol.*, 6: 4295–4304, 1986.
9. Roth, D. B., and Wilson, J. H. Illegitimate recombination in mammalian cells. In: R. Kaucherlapti and G. R. Smith (eds.), *Genetic Recombination*, pp. 621–653. Washington, DC: American Society for Microbiology, 1988.
10. Thode, S., Schafer, A., Pfeiffer, P., and Vielmetter, W. A novel pathway of DNA end-to-end joining. *Cell*, 60: 921–928, 1990.
11. Meuth, M. Illegitimate recombination in mammalian cells. In: D. E. Berg and M. M. Howe (eds.), *Mobile DNA*, pp. 833–860. Washington, DC: American Society for Microbiology, 1989.
12. Reisman, D., Yates, J., and Sugden, B. A putative origin of replication of plasmids derived from Epstein-Barr virus is composed of two *cis*-acting components. *Mol. Cell. Biol.*, 5: 1822–1832, 1985.
13. Konopka, A. K. Compilation of DNA strand exchange sites for nonhomologous recombination in somatic cells. *Nucleic Acids Res.*, 16: 1739–1758, 1988.
14. Jankelevich, S., Kolman, J. L., Bodnar, J. W., and Miller, G. A nuclear matrix attachment region organizes the Epstein-Barr viral plasmid in Raji cells into a single DNA domain. *EMBO J.*, 11: 1165–1176, 1992.
15. Schuler, W., Weiler, A., Schuler, A., Phillips, R., Rosenberg, J., Mak, T., Kearney, J., Perry, R., and Bosma, M. Rearrangement of antigen receptor genes is defective in mice with severe combined immunodeficiency. *Cell*, 46: 963–972, 1986.
16. Biedermann, K. A., Sun, J. R., Giaccia, A. J., Tosto, L. M., and Brown, J. M. *scid* mutation in mice confers hypersensitivity to ionizing radiation and a deficiency in DNA double-strand break repair. *Proc. Natl. Acad. Sci. USA*, 88: 1394–1397, 1991.
17. Nalbantoglu, J., Hartley, D., Phear, G., Tear, G., and Meuth, M. Spontaneous deletion formation at the *aprt* locus of hamster cells: the presence of short sequence homologies and dyad symmetries at deletion termini. *EMBO J.*, 5: 1199–1204, 1986.
18. Duesberg, P. H. Cancer genes: rare recombinants instead of activated oncogenes. *Proc. Natl. Acad. Sci. USA*, 84: 2117–2124, 1987.
19. Roth, D. B., and Wilson, J. H. Relative rates of homologous and nonhomologous recombination in transfected DNA. *Proc. Natl. Acad. Sci. USA*, 82: 3355–3359, 1985.
20. Pfeiffer, P., and Vielmetter, W. Joining of nonhomologous DNA double strand breaks *in vitro*. *Nucleic Acids Res.*, 16: 907–924, 1988.
21. Landau, N. R., Schatz, D. G., Rosa, M., and Baltimore, D. Increased frequency of N-region insertion in a murine pre-B cell line injected with a terminal dideoxynucleotidyl transferase retroviral expression vector. *Mol. Cell. Biol.*, 7: 3237–3243, 1987.