Nucleic Acids Research, 1994, Vol. 22, No. 25 5649–5657

Repair of a specific double-strand break generated within a mammalian chromosome by yeast endonuclease I-*Sce*I

Tamas Lukacsovich, Di Yang and Alan S.Waldman*
Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA

Received August 29, 1994; Revised and Accepted November 4, 1994

ABSTRACT

We established a mouse Ltk$^-$ cell line that contains within its genome a herpes simplex virus thymidine kinase gene (*tk*) that had been disrupted by the insertion of the recognition sequence for yeast endonuclease I-*Sce*I. The artificially introduced 18 bp I-*Sce*I recognition sequence was likely a unique sequence in the genome of the mouse cell line. To assess whether an induced double-strand break (DSB) in the genomic *tk* gene would be repaired preferentially by gene targeting or non-homologous recombination, we electroporated the mouse cell line with endonuclease I-*Sce*I alone, one of two different gene targeting constructs alone, or with I-*Sce*I in conjunction with each of the two targeting constructs. Each targeting construct was, in principle, capable of correcting the defective genomic *tk* sequence via homologous recombination. *tk*$^+$ colonies were recovered following electroporation of cells with I-*Sce*I in the presence or absence of a targeting construct. Through the detection of small deletions at the I-*Sce*I recognition sequence in the mouse genome, we present evidence that a specific DSB can be introduced into the genome of a living mammalian cell by yeast endonuclease I-*Sce*I. We further report that a DSB in the genome of a mouse Ltk$^-$ cell is repaired preferentially by non-homologous end-joining rather than by targeted homologous recombination with an exogenous donor sequence. The potential utility of this system is discussed.

INTRODUCTION

The genome of a mammalian cell is likely to experience a spectrum of DNA damage every day. One type of DNA lesion that virtually all organisms must contend with is the formation of double-strand breaks (DSBs). It has been estimated that a mammalian genome suffers about eight DSBs per day (1). Unrepaired DSBs can be lethal or contribute to the induction of abnormal chromosomal rearrangements in eukaryotes (2–7). It has recently been demonstrated that even a single unrepaired DSB in a dispensible plasmid can serve as a signal for cell death in yeast (8), underscoring the importance of mechanisms for efficient repair of DSBs.

Eukaryotic cells potentially have several mechanisms for repairing a DSB. One possible repair pathway is through recombination between the broken sequence and an intact homologous sequence. It is well documented that DSBs can stimulate extrachromosomal homologous recombination in yeast (reviewed in 9) and mammalian cells (reviewed in 10). Several reports on recombination in mammalian cells suggest that a DSB repair mechanism similar to that originally described for recombination in yeast (11) also operates in extrachromosomal recombination of DNA sequences transfected into mammalian cells (10), as well as in gene targeting (12–14). DSBs and single-strand breaks have been shown to stimulate homologous recombination within the yeast genome (9), but there has been no report directly assessing what effect(s) a break introduced into a mammalian genome may have on homologous recombination. Some mammalian cell mutants that are defective in the repair of DSBs have been shown to display elevated levels of sister chromatid exchanges (SCE) when treated with DNA damaging agents (reviewed in 15), although the precise relationship between SCE and homologous recombination remains unclear. Treatment of mouse cells with a chemical inhibitor of DNA break repair has been shown to induce a 4-fold stimulation of intrachromosomal homologous recombination (16). However, in these latter studies it was not clear if chromosomal breaks served as recombinogenic lesions, or whether recombination was stimulated by an indirect mechanism. Understanding how DSBs may influence homologous recombination in mammalian chromosomes can contribute to our understanding of naturally occurring genetic rearrangements, as well as have potential application in designing strategies for improving the efficiency of gene targeting.

Another potential pathway for the repair of a DSB in a mammalian cell is through DNA end-joining, or non-homologous recombination. It has been shown that mammalian cells can efficiently ligate together virtually any two non-cognate ends of DNA (reviewed in 17). The molecular mechanisms for DNA end-joining in mammalian cells have not yet been elucidated; several investigators have recently initiated biochemical characterizations of end-joining activities isolated from

*To whom correspondence should be addressed

mammalian cells (18–22). Efficient end-joining of DNA might be responsible for the efficiency with which cultured mammalian cells randomly integrate transfected DNA molecules into their genomes (so-called 'illegitimate recombination') (17).

To contribute to our understanding of how mammalian cells process DSBs, we wanted to develop an experimental system that would allow us to introduce a specific DSB within the genome of a living cell. We were particularly interested in determining if a genomic DSB might be repaired efficiently by targeted homologous recombination with a transfected DNA sequence (and therefore be of potential use in the development of a methodology for enhancing gene targeting efficiency) or if, instead, a genomic DSB would be repaired preferentially by non-homologous joining of the broken chromosomal ends.

In our experimental system, we wanted to be able to introduce a single break at a defined genetic locus. Wholesale genomic breakage, as occurs in methods involving the introduction of bacterial restriction enzymes into mammalian cells (see 23 for example), may induce cellular responses to global damage and this may obscure the effect of any single specific DSB. (Massive genomic breakage is also clearly undesirable in gene targeting strategies.) To accomplish our goals, we made use of the endonuclease I-*Sce*I which is encoded by a mobile type I intron from the mitochondrial genome of *Saccharomyces cerevisiae* (24). The recognition sequence for I-*Sce*I cleavage is 18 bp in length, with some tolerance for single bp degeneracies (24). Arithmetical calculations reveal that the I-*Sce*I recognition site is not likely to occur randomly in a mammalian genome.

We designed a DNA construct containing a herpes simplex virus type one (HSV-1) thymidine kinase gene (*tk*) disrupted by the insertion of an 18 bp oligonucleotide containing the I-*Sce*I recognition site. This disrupted *tk* gene was introduced into the genome of mouse Ltk$^-$ (thymidine kinase deficient) fibroblasts, establishing a unique, or nearly unique, I-*Sce*I target sequence within the mouse genome. We reasoned that subsequent introduction of endonuclease I-*Sce*I into the cell could, in principle, lead to the reconstruction of a functional *tk* gene via the induction of a DSB at the I-*Sce*I recognition site followed by appropriate resection and joining of DNA ends to restore a coding sequence. Alternatively, if an appropriate targeting construct were introduced in conjunction with I-*Sce*I, a DSB induced by I-*Sce*I might be repaired via homologous recombination with the targeting construct.

In this work, we show that an I-*Sce*I recognition site embedded within the genome of mouse fibroblasts can be cleaved by electroporating endonuclease I-*Sce*I into the cells. This is the first reported use of an endonuclease to induce a specific DSB at a single site within the genome of a mammalian cell. We further show that such an induced genomic DSB is repaired preferentially by resection and end-joining, rather than by homologous recombination with an exogenous donor sequence.

MATERIALS AND METHODS

Materials

Meganuclease I-*Sce*I was purchased from Boehringer Mannheim (Indianapolis) and was used *in vitro* as prescribed by the supplier.

Cell culture

Mouse Ltk$^-$ cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum, 0.1 mM MEM non-essential amino acids (GIBCO), and 50 μg/ml gentamicin sulfate. To select for *tk*$^+$ cells, medium supplemented with hypoxanthine/aminopterin/thymidine (HAT) was prepared using 100× HAT supplement (GIBCO). Cells were maintained at 37°C in a humidified atmosphere of 5% CO_2.

Plasmid constructions

All plasmids used are based on vector pJS-1, which is a derivative of pSV2neo (25) with restriction site modifications as previously described (26). Plasmid pTK1 (8.3 kb) contains the wild-type HSV-1 *tk* gene on a 2.5 kb *Bam*HI fragment inserted into the unique *Bam*HI site of the vector pJS-1. Into the *Sac*I site at position 960 of the *tk* coding region (numbering according to 27) was inserted the following double-stranded oligonucleotide:

```
                    ↓
      5'-T A G G G A T A A C A G G G T A A T AGCT-3'
3'-TCGAA T C C C T A T T G T C C C A T T A -5'
                    ↑
```

This oligonucleotide contains the recognition sequence for endonuclease I-*Sce*I (24), shown in bold. This oligonucleotide also terminates in sticky ends compatible with ends produced by *Sac*I cleavage. Endonuclease I-*Sce*I cleaves to the 3′ side of the nucleotides indicated by arrows, leaving sticky ends with 3′ single-strand extensions. Insertion of the oligonucleotide at the *Sac*I site of the *tk* gene inactivated the *tk* gene by introducing a frame-shift (a net gain of 22 nucleotides was caused by the insertion, when the duplicated 4 bp terminal sticky ends are taken into account). The plasmid containing the *tk* gene disrupted by the I-*Sce*I recognition site was named pTK1-Sce (Fig. 1). Plasmid pTARG is identical to pTK1, except that the *tk* gene is rendered non-functional due to a 104 bp inversion of the sequences between the two *Eco*RV sites at positions 744 and 848 of the *tk* coding region (Fig. 1). Plasmid pTARG has been described previously (28). A similar plasmid, pAL2 (Fig. 1), contains a defective internal fragment of the *tk* gene inserted at the *Hin*dIII site of the vector. This *tk* fragment maps between the *Hin*cII and *Sma*I sites at positions 442 and 1621 of the *tk* gene. Plasmid pAL2 (7.0 kb) has been described previously (29).

Assay for transient *tk* expression

To establish conditions for efficient electroporation of mouse Ltk$^-$ cells, we introduced pTK1 (described above) into mouse Ltk$^-$ cells under a variety of conditions and assayed the electroporated cells for transient *tk* gene expression. The assay for a functional *tk* gene provided a minimal estimate for the percentage of cells that had taken up DNA. Following electroporation, cells were plated into 35 mm dishes at a density of 1.5×10^5 cells per dish. Five hours later, medium was changed to medium supplemented with 20 μCi/ml [methyl-^{3}H]thymidine (80 Ci/mmol). Cells were allowed to incubate for 24 h in the presence of [methyl-^{3}H]thymidine, were washed with PBS, fixed with methanol, and coated with autoradiography emulsion. Following development of the emulsion, cells were viewed under phase-contrast microscopy. A cell was scored positive for *tk* if it contained heavily localized grains over the nucleus. This *tk* assay was described in greater detail previously (30).

Isolation of cell line Sce-3

5×10^6 mouse Ltk$^-$ cells were resuspended in 800 μl of phosphate-buffered saline (PBS) containing 4 μg pTK1-Sce that had been linearized with *Cla*I. Cells were electroporated at room temperature in a Bio-Rad Gene Pulser using a setting of 1.2 mV

and 25 μF. Cuvettes with a 0.4 cm electrode gap were used. Cells were then plated into 100 mm dishes at a density of $\sim 5 \times 10^4$ viable cells/dish. Two days later the medium was changed to medium supplemented with 400 μg/ml (total drug) G418 (GIBCO). Colonies were harvested 12 days later. DNA samples isolated from individual clones were analyzed on Southern blots using appropriate restriction digestions and an HSV-1 *tk*-specific probe to identify cell lines containing a single integrated copy of pTK1-Sce. Such cell lines were tested for the stability of integration of pTK1-Sce as follows. Ten individual cells from a given cell line were propagated as separate subclones for greater than 10 generations under no genetic selection. Each of the 10 subclones was then tested for resistance to G418. A parental cell line was considered stable if and only if 10 out of 10 subclones retained resistance to G418. One such cell line, designated Sce-3, was used in further studies.

Electroporation of endonuclease I-*Sce*I and DNA targeting constructs into cells

Electroporation conditions were carried out essentially as above except that 10×10^6 mouse cells (cell line Sce-3, described above) were used per electroporation in a volume of 800 μl of PBS. Cells were electroporated with either 150 units of I-*Sce*I alone, or 10 μg (*Cla*I-linearized) pTARG or pAL2 alone, or a mixture of I-*Sce*I plus pTARG or pAL2. In electroporations involving I-*Sce*I, undiluted endonuclease I-*Sce*I in 15 μl storage buffer (as supplied by Boehringer Mannheim) was added directly to the electroporation cuvette just prior to electroporation. Cells suspended in room temperature PBS (plus or minus plasmid DNA) were added to the cuvette, mixed with the endonuclease, and electroporated.

Following electroporation, cells were immediately diluted into room temperature DMEM and plated at a density of 5×10^6 cells (viable plus non-viable cells) per 150 cm^2 flask. Two days later, the medium was changed to HAT medium to select for tk^+ clones. Colonies were counted after 14–18 days growth in HAT medium, with no refeeding.

In experiments to determine cell survival, 100 cells were plated into DMEM in each of three 25 cm^2 flasks immediately following electroporation. Approximately 1 week later, cell survival was assessed in terms of the average number of colonies per flask.

Genomic DNA isolation and Southern blotting analysis

Genomic DNA was prepared from cultured cells and analyzed by Southern hybridization using a ^{32}P-labeled probe specific for the HSV-1 *tk* sequence as described (29).

PCR amplification

HSV-1 *tk* sequences were PCR-amplified from genomic DNA isolated from HAT-resistant (HAT^r) derivatives of mouse cell line Sce-3 using the following set of primers: 5′-TCTACACC-ACACAACACCGC-3′ and 5′-ACAAACGACCCAACA-CCCGT-3′. (Primers were synthesized by the Biochemistry Biotechnology Facility at the Indiana University School of Medicine, Indianapolis.) The first primer is the sequence of the non-coding strand of the HSV-1 *tk* gene from nucleotide position 814 to 833 (27), while the second primer is the HSV-1 *tk* coding sequence from nucleotide position 1734 to 1715 (27). PCR amplifications were carried out in a 100 μl reaction cockail containing: 333 ng mouse genomic DNA; 1.5 μM of each DNA primer; 10 mM Tris–HCl, pH 8.3; 50 mM KCl; 1.1 mM $MgCl_2$; 200 μM of each of the four deoxynucleotide triphosphates; 2.5 units AmpliTaqTM DNA polymerase (Perkin Elmer Cetus, Norwalk, CT). Each reaction cycle consisted of 1 min at 95°C, 1 min at 62°C, 2 min at 72°C. PCR amplification was accomplished by 35 reaction cycles.

DNA sequence determination

PCR products (920 bp in length) were purified by phenol extraction and ethanol precipitation. Purified PCR products were cleaved with *Sph*I and *Msc*I, cloned into M13mp18, and sequenced from single-stranded DNA templates using a Sequenase Version 2.0 DNA Sequencing Kit (US Biochemical).

RESULTS

Experimental design

Our goal was to assess whether a chromosomal DSB in a mammalian cell would more likely be repaired via non-homologous DNA end-joining or via targeted homologous recombination. We reasoned that if homologous recombination were indeed the preferred route for repair of a genomic DSB (and if recombinational repair is actually *induced* in the vicinity of a genomic DSB), then strategies aimed at targeted induction of a genomic DSB might ultimately prove useful in developing methodologies for enhancing gene targeting efficiency. As a step toward accomplishing our more immediate goal, we wanted to engineer a mouse cell line to contain a unique chromosomal sequence that could serve as a target for specific endonucleolytic cleavage. This target sequence would also serve as a defective selectable marker that could be corrected by gene targeting or, if the sequence were broken, by certain non-homologous end-joining events. The engineered cell line would be electroporated with a specific endonuclease to induce a DSB in the genomic target sequence in the presence or absence of appropriate gene targeting constructs. A determination would be made of whether reconstruction of the broken marker locus occurred preferentially via targeted homologous recombination or non-homologous end-joining.

In order for an induced genomic DSB to be repaired via targeted homologous recombination, it is clearly imperative that exogenous homologous donor DNA sequences be present in the cell, i.e. targeting DNA constructs must be delivered efficiently into the population of cells in which DSBs are induced. Using an assay for transient *tk* expression, we established electroporation conditions under which we can detect *tk* gene expression in 29% of a population of mouse Ltk^- cells electroporated with a plasmid containing a wild-type HSV-1 *tk* gene. By using the established electroporation conditions, we could be reasonably confident that if an endonuclease were co-electroporated with a targeting construct into mouse Ltk^- cells, one or more copies of the DNA construct would be introduced into at least 29% of the cells which took up the endonuclease.

As described above (see Introduction), it was important in our experimental strategy that the endonuclease recognition sequence represent a unique (or nearly unique) genomic sequence. We decided to introduce the 18 bp recognition sequence (5′-TAGGG-ATAACAGGGTAAT-3′) for yeast endonuclease I-*Sce*I into the genome of mouse Ltk^- cells, because straightforward arithmetical calculations predict that any specific 18 bp sequence will randomly occur in nature less than once per 10^{10} bp of DNA complexity. We predicted that the 18 bp I-*Sce*I recognition sequence is thus not likely to be present in a typical mammalian

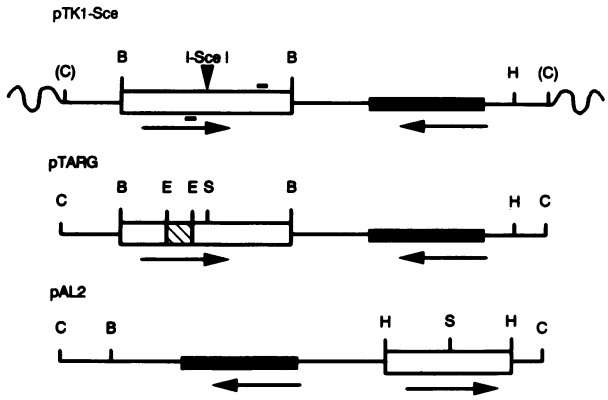

Figure 1. Structure of DNA constructs. Constructs are illustrated in linearized form, cleaved at the unique *ClaI* site in the vector. Inserts containing HSV-1 *tk* sequences are depicted by open rectangles, the *neo* gene is depicted by filled rectangles. The 104 bp *tk* sequence inversion in pTARG is shown as a hatched segment. Arrows indicate transcription direction. Restriction sites are: C, *ClaI*; B, *BamHI*; H, *HindIII*; E, *EcoRI*; S, *SacI*. Also indicated is the I-*SceI* recognition site inserted in the *tk* gene in pTK1-Sce. Construct pTK1-Sce is depicted as it resides within the genome of mouse L cell line Sce-3, with flanking genomic sequences drawn as wavy lines. The *ClaI* site of pTKI-Sce was likely lost upon integration of pTK1-Sce into the mouse genome. Short bars above and below the *tk* sequence of pTK1-Sce indicate the positions of primers used for PCR amplification. Figure not to scale.

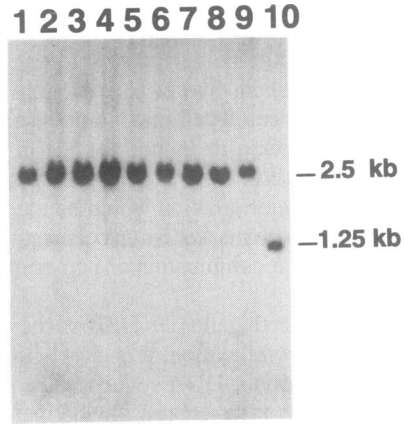

Figure 2. Disposition of the I-*SceI* recognition site in cell line Sce-3 and HAT^r derivatives of cell line Sce-3. Genomic DNA (8 μg) isolated from parental cell line Sce-3 was digested with *BamHI* (lane 9) or *BamHI* plus I-*SceI* (lane 10) and displayed on a Southern blot using a probe specific for the HSV-1 *tk* gene sequence. The 2.5 kb *BamHI* fragment containing the HSV-1 *tk* (lane 9) was cleaved by I-*SceI* into two 1.25 kb fragments (lane 10), demonstrating the presence of a functional I-*SceI* recognition sequence in the genomic HSV-1 *tk* gene of cell line Sce-3. DNA samples (8 μg) isolated from eight HAT^r clones (generated by electroporating line Sce-3 with I-*SceI* plus pTARG) were digested with *BamHI* plus I-*SceI* and are displayed in lanes 1–8. Each HAT^r clone contains a 2.5 kb *BamHI* fragment containing an HSV-1 *tk* gene that is not cleavable by I-*SceI*, indicating that the I-*SceI* recognition sequence had been altered in each of the HAT^r derivatives of cell line Sce-3.

genome. Consistent with this expectation, when genomic DNA isolated from mouse L cells was incubated with I-*SceI* and displayed on an agarose gel using ethidium bromide staining, we observed no detectable cleavage compared to untreated mouse

Table 1. Recovery of HAT^r colonies following electroporation of endonuclease I-*SceI* into mouse cell line Sce-3

Electroporated molecules[a]	No. of electroporations[b]	No. of HAT^r colonies
I-*Sce* I	9 (4)	5
pTARG	6	0
I-*Sce* I & pTARG	6 (4)	10
pAL2	9	0
I-*Sce* I & pAL2	12 (2)	8

[a]For each electroporation, 10×10^6 cells were electroporated with one of the following: 150 units endonuclease I-*SceI*, 10 μg pTARG or pAL2 DNA, or a mixture of 150 units I-*Sce* I plus 10 μg either pTARG or pAL2 DNA.
[b]Numbers in parentheses indicate the number of electroporations that yielded at least one HAT^r colony.

L cell genomic DNA (data not shown). In addition, a computer search of GenBank (release 80.0, 12/93) and EMBL (release 36.0, 9/93) databases using the Wisconsin GCG software package (version 7.3) did not reveal any higher eukaryotic sequence that matched either 16, 17, or the complete 18 bp of the I-*SceI* recognition sequence. Based on these calculations and observations, we concluded that an I-*SceI* recognition sequence artificially introduced into the mouse L cell genome would likely represent a unique, or nearly unique, sequence.

Establishing a cell line containing a genomic I-*SceI* target sequence

We constructed plasmid pTK1-Sce (Fig. 1). This plasmid contains an HSV-1 *tk* gene with a coding region that is disrupted by an insertion of an oligonucleotide containing the 18 bp I-*SceI* recognition sequence. We confirmed by DNA sequence analysis that the *tk* gene in pTK1-Sce contained only a single copy of the I-*SceI* recognition sequence. (The left-to-right orientation, with respect to Fig. 1, of the I-*SceI* recognition sequence in pTK1-Sce is the same as the orientation presented at the top of Fig. 4.) We also confirmed by agarose gel electrophoresis that supercoiled molecules of pTK1-Sce could be cleaved with I-*SceI* *in vitro* (not shown).

Plasmid pTK1-Sce was linearized with *ClaI* (as depicted in Fig. 1) and was electroporated into mouse Ltk$^-$ cells to establish G418-resistant cell lines containing one or several copies of the construct stably integrated in the mouse L cell genome. One such cell line, designated Sce-3, which contained only a single integrated copy of pTK1-Sce, was used in further studies.

To confirm that the genome of cell line Sce-3 harbored a *tk* gene containing a cleavable copy of the I-*SceI* recognition sequence, genomic DNA isolated from Sce-3 cells was digested with *BamHI* or *BamHI* plus I-*SceI*. Southern blotting analysis using a probe specific for the HSV-1 *tk* gene revealed a 2.5 kb *BamHI* fragment that could be cleaved into two 1.25 kb fragments by I-*SceI* (Fig. 2, lanes 9 and 10). This analysis thereby corroborated that the genome of Sce-3 cells contained a *tk* gene that was cleavable, *in vitro*, with I-*SceI*.

Recovery of HAT^r derivatives of cell line Sce-3 following introduction of endonuclease I-*SceI* via electroporation

We intended to electroporate cell line Sce-3 with endonuclease I-*SceI*, pTARG, and pAL2 (Fig. 1) in various combinations and select for HAT^r cells. Before conducting these studies, we first assessed whether the introduction of endonuclease I-*SceI* into mouse fibroblasts would substantially reduce cell viability. In a representative experiment involving electroporation of mouse

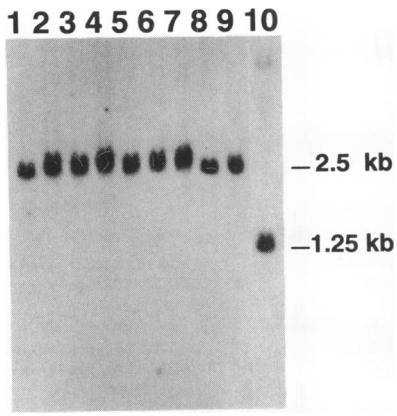

Figure 3. HAT^r derivatives of cell line Sce-3 do not contain a wild-type *tk* sequence. DNA samples (8 μg) isolated from the eight HAT^r clones exhibited in Fig. 2 were digested with *Bam*HI plus *Sac*I and displayed on a Southern blot (lanes 1−8) using a probe specific for the HSV-1 *tk* gene. Each HAT^r clone contains a 2.5 kb *Bam*HI fragment that cannot be cleaved by *Sac*I into two 1.25 kb fragments, indicating that a wild-type *tk* sequence (which contains a *Sac*I recognition site) is not present in any of the HAT^r clones. Also displayed as size markers are *Bam*HI (lane 9) and *Bam*HI plus I-*Sce*I (lane 10) digests of DNA from parental cell line Sce-3.

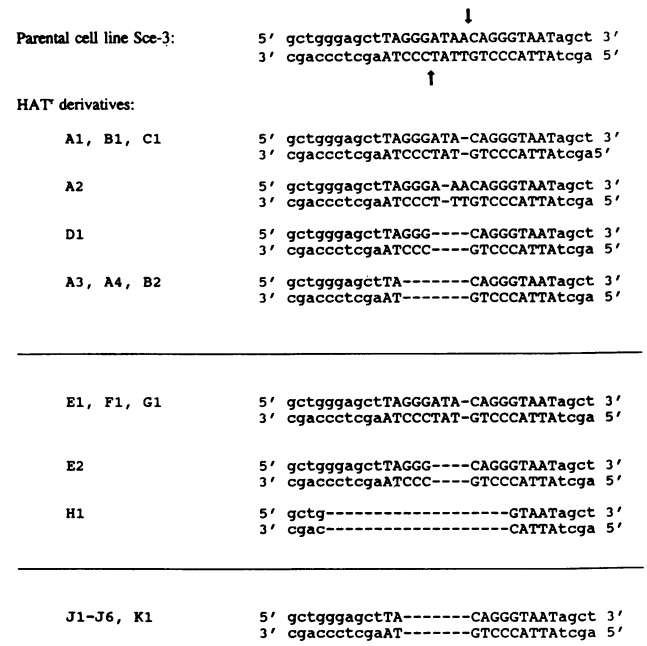

Figure 4. Nucleotide sequences from HAT^r derivatives of cell line Sce-3. Shown are the sequences in the immediate vicinity of the I-*Sce*I recognition site in parental cell line Sce-3, as well as the corresponding sequences from HAT^r derivatives. The I-*Sce*I recognition sequence is denoted by uppercase characters. Arrows indicate the predicted sites of cleavage by endonuclease I-*Sce*I in the top and bottom strands of the parental sequence. The HAT^r clones are grouped according to the type of electroporation that generated the HAT^r clones; clones whose names begin with the same letter were isolated from the same flask of cells following electroporation. HAT^r clones were generated by electroporating cell line Sce-3 with endonuclease I-*Sce*I plus pTARG (clones A1−D1, upper section of figure), I-*Sce*I alone (clones E1−H1, middle section), or I-*Sce*I plus pAL2 (clones J1−K1, bottom section). Deleted nucleotides in sequences from the HAT^r derivatives are denoted by dashes. There is some ambiguity in the precise endpoints of the deletion in clone H1, because the deletion in this clone terminates in two stretches of three consecutive G residues in the parental sequence.

Ltk$^-$ cells with PBS alone or PBS plus 150 units of I-*Sce*I, plating efficiencies following electroporation were 85% and 74%, respectively. Additionally, cell viability was the same whether cells were electroporated with 10 μg linearized pTARG DNA alone or with a mixture of 10 μg pTARG DNA plus 150 units I-*Sce*I. From such determinations, we concluded that electroporation of I-*Sce*I into mouse Ltk$^-$ cells had little or no effect on cell viability. This contrasts with the high degree of lethality (plating efficiencies below 10%) reported for some experiments involving electroporation of mammalian cells with restriction enzymes having 4 bp recognition sites (23).

Sce-3 cells were electroporated with I-*Sce*I alone, pTARG or pAL2 alone, or I-*Sce*I in combination with pTARG or pAL2, as indicated in Table 1. Two days following electroporation, cells were fed with HAT medium and incubated for an additional 2 weeks to select for *tk*$^+$ cells. The numbers of HAT^r colonies recovered following the various electroporations are listed in Table 1. HAT^r colonies were recovered following only the electroporations involving endonuclease I-*Sce*I, irrespective of the presence or absence of pTARG or pAL2 (Table 1).

Analysis of genomic *tk* sequences from HAT^r derivatives of cell line Sce-3

We inferred that HAT^r clones that arose after electroporations were due to the reconstruction of a functional HSV-1 *tk* in the genome of Sce-3 cells. In electroporations involving pTARG or pAL2, *tk* gene reconstruction may have occurred through gene targeting, since pTARG and pAL2 (Fig. 1) each contain sufficient genetic information to potentially correct the defective genomic *tk* gene in cell line Sce-3 via homologous recombination. (It seemed unlikely that the defective *tk* gene in pTARG or pAL2 would be corrected by recombination with the genomic *tk* gene because of the greater severity of the mutation in the pTARG or pAL2 *tk* sequence compared with the mutation in the genomic *tk* sequence.) Alternatively, *tk*$^+$ clones could have been produced via resection of DNA termini from a DSB induced by I-*Sce*I followed by ligation of the DNA ends to restore the *tk* reading frame.

To analyze the events responsible for generating the HAT^r colonies, HAT^r clones were expanded and genomic DNA isolated. We initially focused our analysis on eight clones that arose following electroporation of cell line Sce-3 with 150 units I-*Sce*I plus 10 μg pTARG. Southern blotting analysis was performed (Fig. 2) to ascertain whether the I-*Sce*I recognition sequence in the genomic HSV-1 *tk* sequence had been altered in each of the clones. Digestion of DNA samples with *Bam*HI in combination with I-*Sce*I revealed that the I-*Sce*I recognition sequence in the *tk* gene of each HAT^r clone had been altered, since each clone contained a *tk* gene that was no longer cleavable with I-*Sce*I. (Compare HAT^r clones, Fig. 2, lanes 1−8, with parental line Sce-3, Fig. 2, lanes 9 and 10.)

The I-*Sce*I recognition sequence inserted in the HSV-1 *tk* coding sequence in parental cell line Sce-3 disrupted a *Sac*I site normally present in the wild-type *tk* sequence. This *Sac*I site was contained in the *tk* sequence of pTARG (Fig. 1). Restoration of the genomic HSV *tk* sequence to a wild-type sequence via targeted homologous recombination with pTARG would therefore have restored the *Sac*I site to the genomic *tk* sequence. As displayed

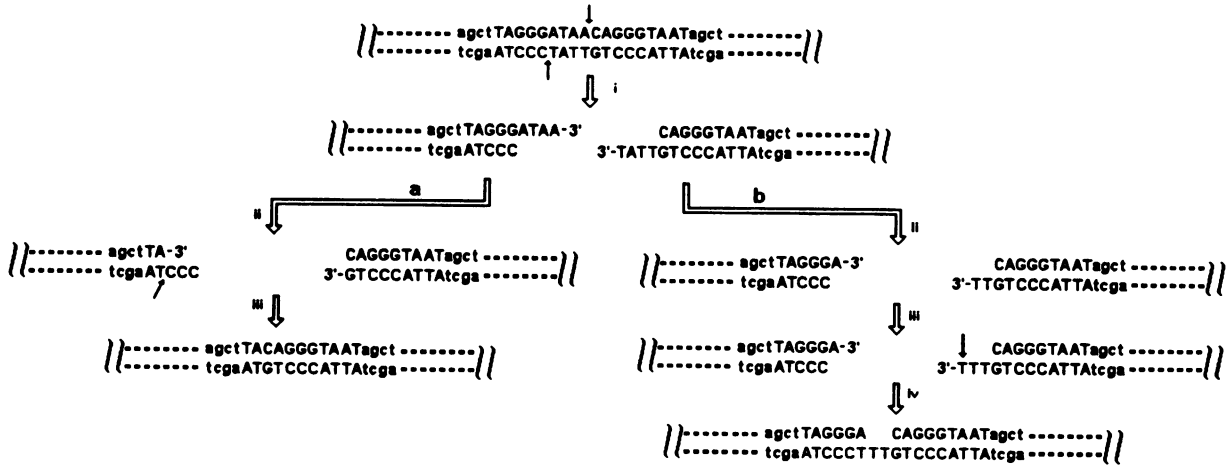

Figure 5. Possible mechanisms for generating the small genomic deletions in HAT^r derivatives of cell line Sce-3. Deletion formation begins with the introduction of a staggered break by endonuclease I-*Sce*I (step i). Formation of sequences with 7 bp deletions (Fig. 4) may have proceeded through pathway a: (ii) loss of terminal nucleotides through 3′→5′ exonucleolytic digestion; (iii) loss of 5′ single-stranded tail on the bottom strand, ligation of strands. Other HAT^r clones with smaller deletions, with the exception of clone A2 (Fig. 4), may have arisen through pathways similar to pathway a. Clone A2 may have been formed via pathway b: (ii) loss of several 3′ terminal nucleotides; (iii) addition of filler nucleotide; (iv) terminal pairing followed by ligation and gap-filling. See text for details.

by a Southern blot (Fig. 3), none of the eight HAT^r clones analyzed contained a *tk* sequence with a restored *Sac*I site. This indicated that none of the HAT^r clones arose from accurate homologous recombination with pTARG. Additionally, none of the HAT^r clones contained any randomly integrated copies of pTARG.

Seven HAT^r clones recovered after electroporation of parental cell line Sce-3 with I-*Sce*I plus pAL2 (Table 1) were analyzed by Southern blotting of genomic sequences following digestion with *Bam*HI plus I-*Sce*I, or *Bam*HI plus *Sac*I. This analysis revealed that the *tk* genes in these clones, like the eight other clones discussed above, could not be cleaved with either I-*Sce*I or *Sac*I (data not shown). None of these clones contained additional HSV-1 *tk* sequences compared with the parental cell line. These results again suggested that resection and ligation of DNA termini, rather than homologous recombination with targeting construct pAL2, was responsible for the repair of the DSB induced by I-*Sce*I.

PCR amplification was used to isolate a segment of the genomic *tk* sequence from 20 HAT^r clones that arose after electroporating cell line Sce-3 with endonuclease I-*Sce*I alone or in combination with pTARG or pAL2. Approximately 200 bp of nucleotide sequence spanning the position of the I-*Sce*I recognition site was determined for each clone. The relevant sequences are shown in Fig. 4. HAT^r clones A1−D1 (top section of Fig. 4), which arose following electroporation of cell line Sce-3 with I-*Sce*I plus pTARG, are the very same clones displayed on the Southern blots in Figs. 2 and 3. HAT^r clones E1−H1 (middle section of Fig. 4) arose after electroporation of cell line Sce-3 with I-*Sce*I alone, while clones J1−J6 plus K1 (bottom section of Fig. 4) arose after electroporation of cell line Sce-3 with I-*Sce*I plus pAL2.

Compared with the *tk* sequence of parental line Sce-3, the only sequence alterations noted in all but one of the HAT^r derivatives of Sce-3 resided entirely within the I-*Sce*I recognition sequence itself. Clone H1 (Fig. 4) contained a 19 bp deletion that penetrated several nucleotides into sequences flanking the I-*Sce*I recognition sequence. All remaining HAT^r clones displayed a 1, 4, or 7 bp deletion in the immediate vicinity of the predicted I-*Sce*I cleavage site. All deletions restored the reading frame of the *tk* gene, since the original insertion of the I-*Sce*I recognition sequence was equivalent to a 22 bp insertion. Additionally, inspection of the nucleotide sequence of each reverted *tk* gene revealed no in-frame stop codons.

The C residue immediately to the right of the I-*Sce*I cleavage site in the top strand, as depicted at the top of Fig. 4, was preserved in 19 of 20 HAT^r clones and bordered the deletion in 18 of the clones. The C residue immediately to the left of the cleavage site in the bottom strand was preserved in nine out of 20 sequences. The sequence data collectively suggested that, for each HAT^r clone recovered, a DSB had been introduced by endonuclease I-*Sce*I within the recognition sequence in the genome of the parental line Sce-3 *tk* gene. The DNA termini produced at the I-*Sce*I break site were apparently processed by removal of one or a few nucleotides and the DNA ends ligated back together to restore *tk* gene function (see Discussion). Resection and ligation appeared to be more efficient than targeted homologous recombination with pTARG or pAL2 as a means for repairing the chromosomal DSB induced by I-*Sce*I.

DISCUSSION

In this report, we have presented evidence that a DSB may be introduced within a chromosome of a living mammalian cell by electroporating yeast endonuclease I-*Sce*I into a cell. This is the first reported use of an endonuclease to induce a specific DSB at a single site within the genome of a mammalian cell. The ability to introduce a DSB at a unique, or nearly unique, site in the genome of a living mammalian cell provides a potentially powerful tool with which to investigate cellular responses to a chromosomal break without interference from effects of more global genomic damage. Our evidence for the induction of a chromosomal DSB by I-*Sce*I consists of the recovery of *tk*$^+$ cells due to deletion formation at the putative break site. (Attempts at directly visualizing the DSB by Southern blotting were unsuccessful.) The following observations were made: (i) HAT^r

(tk^+) cells were recovered following only electroporations involving endonuclease I-*Sce*I; (ii) compared to the parental cell line Sce-3, none of the recovered HAT^r clones contained any additional copies, or gross-rearrangements, of genomic HSV-1 *tk* sequences; (iii) 20 out of the 20 HAT^r clones recovered and analyzed contained deletions that restored the *tk* gene reading frame and mapped to the immediate vicinity of the predicted cleavage site for I-*Sce*I (Fig. 4). These observations provide strong evidence that electroporated I-*Sce*I did indeed cleave its recognition sequence embedded in the genome of mouse fibroblasts. Because of the imposition of a genetic selection for tk^+ clones only, we could not accurately calculate the fraction of the cells electroporated with I-*Sce*I that actually had a DSB induced in their genomes. It is conceivable that the actual frequency of DSB induction (and repair) significantly exceeded the number of HAT^r colonies recovered. For example, the generation of 4 bp single-stranded sticky ends by cleavage with I-*Sce*I may provide sufficient sequence complementarity to allow efficient annealing of the sticky ends and resealing of the DSB without alteration of the sequence or restoration of *tk* gene function. Deletions that did not restore the reading frame, or that removed codons for critical amino acids, would also not be recovered.

An interesting finding is that repair of the DSB induced by I-*Sce*I occurred preferentially via non-homologous end-joining rather than targeted homologous recombination with pTARG or pAL2. As discussed above, under the electroporation conditions used in this study we can deliver DNA to at least 29% of the electroporated cell population. We may therefore infer that of the 15 HAT^r clones that were analyzed following co-electroporation of I-*Sce*I with pTARG or pAL2 (see Fig. 4), at least 4 or 5 of these clones likely took up one or more copies of a gene-targeting construct along with endonuclease I-*Sce*I. However, none of the HAT^r clones contained a *tk* gene corrected by homologous recombination. Rather, they all displayed small deletions in the vicinity of the I-*Sce*I site, suggestive of a non-homologous end-joining process (see below). We conclude that non-homologous end-joining is more efficient than targeted homologous recombination as a means for repairing a chromosomal DSB in a mammalian cell. We have examined a single locus in a single cell type; it is possible that the relative kinetics of different repair pathways vary from locus to locus and from one cell type to another.

Previous work by others (31,32), as well as ongoing work in our laboratory, indicates that the initial homologous interaction between a transfected DNA molecule and a genomic target is not the rate limiting step of gene targeting in mammalian cells. [In contrast, it was recently reported that the initial homologous interaction is rate limiting for gene targeting in yeast (33).] Additionally, we previously determined that homologous pairing is not the rate limiting step for extrachromosomal recombination in mammalian cells (28). Based on these observations, it seems reasonable to suggest that a step following homologous pairing is rate limiting for gene targeting in mammalian cells. Our recent extrachromosomal recombination studies in mammalian cells (28) indicated that a DSB may help commit a synaptic complex to recombination. In a similar vein, and in consideration of the DSB repair model originally invoked for recombination in yeast (11) and recently invoked as a mechanism for gene targeting in mammalian cells (12–14), we hypothesized that a DSB in a chromosomal sequence would enhance gene targeting at the break site. The experiments reported here were intended as a test of this hypothesis.

We have failed to demonstrate the usefulness of induction of a chromosomal DSB as a strategy for improving gene targeting efficiency in mammalian cells. (The possibility of directing the non-homologous integration of a transfected DNA molecule into a genomic DSB remains unexplored.) The striking inefficiency of targeted homologous recombination in our system (using two different targeting constructs) in the absence of I-*Sce*I makes it difficult to rigorously assess whether induction of a chromosomal DSB has *any* effect on the rate of the targeting process. Nevertheless, our studies suggest that the obstacle to efficient gene targeting in our experimental system using mouse Ltk^- cells cannot be overcome by the placement of a DSB in the target sequence.

The terminal sequences at the genomic DSB induced by I-*Sce*I did not share homology with the targeting constructs pTARG or pAL2 due to the I-*Sce*I recognition sequence that is present in the genomic sequence but absent from the targeting constructs. It is possible that this heterology at the break site inhibited repair of the genomic break via homologous recombination. Gene targeting might be stimulated by a genomic DSB that has terminal nucleotide sequences homologous to the targeting sequences. We are currently designing a suitable targeting construct that has homology to the DNA termini at the genomic break site to test this hypothesis. (It should be noted that removal of 10 or more nucleotides from either side of the I-*Sce*I-induced DSB prior to end-joining, a type of event that we did not recover, would have produced DNA termini sharing homology to the targeting constructs.)

Using several different *in vivo* and *in vitro* systems, others have demonstrated that mammalian cells can efficiently join together virtually any two ends of DNA with minimal loss of terminal nucleotides (18–22,34 and reviewed in 17). These previous studies did not examine repair of chromosomal DSBs. Several studies have addressed restriction endonuclease-inflicted damage to the entire mammalian genome and have implicated direct ligation as a possible mechanism for repair of chromosomal DSBs (reviewed in 35). Examination at the nucleotide level of the genomic deletions produced in our current studies, ostensibly created by cleavage by I-*Sce*I followed by nucleotide loss and subsequent end-joining, reveals some information on the end-joining process as it occurs in mammalian chromosomes. It seemed striking that the analyzed deletions in the vicinity of the I-*Sce*I-induced DSB were very small. Nineteen of 20 deletions were 1–7 bp in length and were contained entirely within the I-*Sce*I recognition site (Fig. 4). Seven of the deletions were only 1 bp in length, making it seem implausible that the small sizes of the deletions were merely a reflection of selection constraints on the types of events that were recoverable. The recovery of only small deletions is consistent with the notion that mammalian cells repair chromosomal DSBs rapidly, before sufficient time elapses to allow substantial nucleotide loss from the termini of the break.

With the exception of HAT^r clone A2 (Fig. 4), the formation of each deleted sequence may be explained by mechanisms similar to the one depicted in Fig. 5a for the formation of the 7 bp deletion in clones A3, A4, B2, J1–J6, and K1 (Fig. 4). Cleavage by I-*Sce*I produces 3′ single-stranded tails, 4 bp in length (Fig. 5, step i). The single-stranded tails may be degraded by a 3′→5′ exonuclease, which may continue beyond the point of producing

blunt ends (Fig. 5a, step ii). Following removal of the 3′ tails, the 5′ terminal C in the top strand may be joined to a juxtaposed 3′ end, the unpaired single-stranded DNA tail (CCC) in the bottom strand removed, and the bottom strand ligated closed (Fig. 5a, step iii). The model presented in Fig. 5a (which is certainly not the exclusive mechanism by which the 7 bp deletions in the genomes of HAT^r clones could have arisen) is nearly identical to one recently proposed by Derbyshire *et al.* (18) for non-homologous recombination in human cells.

One way in which the deleted *tk* sequence in clone A2 (Fig. 4) could have arisen is by incomplete degradation of 3′ single-stranded tails (Fig. 5b, step ii), addition of a T residue to the 3′ terminus on the bottom strand (Fig. 5b, step iii), pairing of the terminal 3′ A on the top strand with the terminal 3′ T on the bottom strand (Fig. 5b, step iv), followed by filling-in of single-strand gaps and ligation. Addition of one or more nucleotides of 'filler DNA' during non-homologous end-joining (as invoked in Fig 5b, step iii) has been observed in other studies of non-homologous end-joining (see 17). Also, several models for non-homologous end-joining in mammalian cells (17,21,34,36), yeast (37,38), and frog oocytes (39) invoke a role for terminal annealing of very short stretches of complementary sequences, as we have done in our model for the generation of the sequence in clone A2. We also note that clone H1 (Fig. 4) contains a 19 bp deletion with endpoints in two runs of three consecutive G residues in the top strand of the parental sequence. It is possible that these two short stretches of homology in the parental sequence played a role in DNA end-joining.

The retention, in 19 of 20 of the deleted sequences examined (Fig. 4), of the 5′ terminal C residue at the I-*Sce*I cut site in the top strand may be explained in at least two different ways. A 3′→5′ exonucleolytic activity may act asymmetrically such that degradation beyond the formation of blunt ends may occur in a leftward direction (on the top strand) only. Alternatively, additional products of end-joining may be produced that do not contain the C residue in question, but these products are not recovered under HAT selection. In any case, the retention of the C residue at a 5′ terminus of an I-*Sce*I cut site in 19 HAT^r clones suggests that exonucleolytic degradation acted preferentially at 3′ termini. Whether this reflects easy accessibility of nucleases to the protruding 3′ tails produced by I-*Sce*I or is due to a predominant 3′→5′ exonucleolytic activity in mouse cells is not clear. However, it has been reported that when various combinations of non-cognate DNA termini with 5′ or 3′ protruding single-strand tails were end-joined in human cells, there was preferential loss of nucleotides from 3′ tails (18).

A recent study (40) examined loss-of-function mutations induced in the *aprt* gene of Chinese hamster ovary (CHO) cells by the action of electroporated restriction enzymes. Many of the restriction enzyme-induced mutations at the *aprt* locus consisted of small deletions (1−36 bp), or small insertions (1−13 bp), and thus show similarities to the mutations we have described in the present study. However, among the recovered restriction enzyme-induced mutations, ≥10% were complex rearrangements consistent with chromosomal translocations, inversions, or large insertions (40). Formation of such rearrangements might require interactions between two or more chromosomal DSBs, as would be produced within the CHO genome by electroporated restriction enzymes (40). In the present study, we could recover only those I-*Sce*I-induced mutations that produced functional *tk* genes. The issue of whether a single genomic DSB can induce complex chromosomal rearrangements therefore awaits further study.

Repair of a DSB, mutagenic or otherwise, can theoretically occur through a templated or a non-templated mechanism (41). Because we recovered HAT^r colonies after electroporation of mouse cells with I-*Sce*I alone (Table 1, Fig. 4) and, therefore, under conditions in which there was theoretically no homologous template, we inferred that repair of an I-*Sce*I-induced break does not require a template. However, it is conceivable that a template was somehow involved in the resection/ligation repair pathway of the genomic break when an homologous sequence (pTARG or pAL2) was available. Sequence analysis of a much larger sampling of I-*Sce*I-induced chromosomal DSBs repaired in the presence or absence of an homologous template may shed light on this issue.

While preparing this manuscript, we learned of a report on the expression of a gene encoding I-*Sce*I in mammalian cells (42). Constitutive or regulated expression of an I-*Sce*I gene in mammalian cells will likely provide a useful system for studying processing of induced DSBs. Introduction of the gene for endonuclease I-*Sce*I has an advantage over electroporation of the endonuclease itself, in that the endonuclease will be present in virtually every cell of a line transformed with the gene. On the other hand, approaches involving electroporation (or other methods) for directly introducing the I-*Sce*I protein into cells ensure that endonuclease activity will exist only transiently intracellularly, thus avoiding any consequences of chronic DNA breakage that may occur during culturing cells containing the I-*Sce*I gene. Additionally, introduction of the I-*Sce*I protein itself, or other endonucleases, allows for the testing of potential gene targeting strategies that do not involve extraneous genetic augmentation of the cell genome.

ACKNOWLEDGEMENTS

The authors thank Barbara C.Waldman and Nawin Mishra for helpful criticisms of the manuscript. The work was supported by NIH grant GM47110 to A.S.W.

REFERENCES

1. Bernstein,C., and Bernstein,H. (1991) Aging, sex, and DNA repair. Academic Press, San Diego, CA.
2. Fairhead,C. and Dujon,B. (1993) *Mol. Gen. Genet.* **240**, 170−180.
3. Game,J.C. and Mortimer,R.K. (1974) *Mutat. Res.* **24**, 281−292.
4. Hartwell,L. (1992) *Cell* **71**, 543−546.
5. Malone,R.E. and Esposito,R.E. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 503−507.
6. Resnick,M.A. and Martin,P. (1976) *Mol. Gen. Genet.* **143**, 119−129.
7. Weiffenbach,B. and Haber,J.E. (1981) *Mol. Cell. Biol.* **1**, 522−534.
8. Bennett, C.B., Lewis,A.L., Baldwin,K.K. and Resnick,M.A. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5613−5617.
9. Petes,T.D., Malone,R.E. and Symington,L.S. (1991) In Broach,J.R., Pringle,J.R. and Jones,E.W. (eds), The molecular and cellular biology of the yeast *Saccharomyces*: genome dynamics, protein synthesis, and energetics. Cold Spring Harbor Laboratory Press, Plainview, NY, pp. 407−521.
10. Bollag,R.J., Waldman,A.S and Liskay,R.M. (1989) *Annu. Rev. Genet.* **23**, 199−225.
11. Szostak,J.W., Orr-Weaver,T.L., Rothstein,R.J. and Stahl,F.W. (1983) *Cell* **33**, 25−35.
12. Deng,C., Thomas,K.R. and Capecchi,M.R. (1993) *Mol. Cell. Biol.* **13**, 2134−2140.
13. Pennington,S.L. and Wilson,J.H. (1991) *Proc. Natl. Acad. Sci. USA.* **88**, 9498−9502.
14. Valancius,V. and Smithies,O. (1991) *Mol. Cell. Biol.* **11**, 4389−4397.

15. Thompson,L. (1988) In Kucherlapati,R. and Smith,G.R. (eds), Genetic recombination. American Society for Microbiology, Washington, DC, pp. 597–620.
16. Waldman,A.S. Waldman,B.C. (1991) *Nucleic Acids Res.* **19**, 5943–5947.
17. Roth,D. and Wilson,J. (1988) In Kucherlapati,R. and Smith.G.R. (eds), Genetic recombination. American Society for Microbiology, Washinton, DC, pp. 621–654.
18. Derbyshire, M. K., Epstein,L.H., Young,C.S.H, Munz,P.L. and Fishel.R. (1994) *Mol. Cell. Biol.* **14**, 156–169.
19. Fairman,M.P., Johnson,A.P. and Thacker,J. (1992) *Nucleic Acids Res.* **20**, 4145–4152.
20. Fishel,R., Derbyshire,M.K., Moore,S.P. and Young,C.S.H. (1991) *Biochimie* **73**, 257–267.
21. Nicolas,A.L. and Young,C.S.H. (1994) *Mol. Cell. Biol.* **14**, 170–180.
22. North,P., Ganesh,A. and Thacker,J. (1990) *Nucleic Acids Res.* **18**, 6205–6210.
23. Winegar,R.A., Phillips,J.W., Youngblom,J.K. and Morgan,W.F. (1989) *Mutat. Res.* **225**, 49–53.
24. Colleaux,L., d'Auriol,L., Galibert,F. and Dujon,B. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 6022–6026.
25. Southern,P.S. and Berg,P. (1982) *J. Mol. Appl. Genet.* **1**, 327–341.
26. Liskay,R.M., Stachelek,J.L. and Letsou,A. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 183–189.
27. Wagner,M.J., Sharp,J.A. and Summers,W.C. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 1441–1445.
28. Waldman,A.S. (1994) *Genetics* **136**, 597–605.
29. Waldman,A.S. and Liskay,R.M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 5340–5344.
30. Waldman,B.C. and Waldman,A.S. (1990) *Nucleic Acids Res.* **18**, 5981–5988.
31. Thomas,K.R., Folger,K.M. and Capecchi,M.R. (1986) *Cell* **44**, 419–428.
32. Zheng,H., and Wilson,J.H. (1990) *Nature* **344**, 170–173.
33. Wilson,J.H., Leung,W.-Y., Bosco,G., Dieu,D. and Haber,J.E. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 177–181.
34. Roth,D.B. and Wilson,J.H. (1986) *Mol. Cell. Biol.* **6**, 4295–5304.
35. Bryant,P.E. and Johnston,P.J. (1993) *Mutat. Res.* **299**, 289–296.
36. Thacker,J., Chalk,J., Ganesh,A. and North,P. (1992) *Nucleic Acids Res.* **20**, 6183–6188.
37. Kramer,K.M., Brock,J.O., Bloom,K., Moore,J.K. and Haber,J.E. (1994) *Mol. Cell Biol.* **14**, 1293–1301.
38. Mezard,C. and Nicolas,A. (1994) *Mol. Cell. Biol.* **14**, 1278–1292.
39. Pfeiffer,P., Thode,S., Hancke,J. and Vielmetter,W. (1994) *Mol. Cell. Biol.* **14**, 888–895.
40. Phillips,J.W. and Morgan,W.F. (1994) *Mol. Cell. Biol.* **14**, 5794–5803.
41. Harris,R.S., Longerich,S. and Rosenberg,S.M. (1994) *Science* **264**, 258–260.
42. Rouet,P., Smih,F. and Jasin,M. (1994) *Proc. Natl. Acad. Sci USA* **91**, 6064–6068.