Introduction of Double-Strand Breaks into the Genome of Mouse Cells by Expression of a Rare-Cutting Endonuclease

PHILIPPE ROUET, FATIMA SMIH, AND MARIA JASIN*

Cell Biology and Genetics Program, Sloan-Kettering Institute and Cornell University
Graduate School of Medical Sciences, New York, New York 10021

Received 28 June 1994/Returned for modification 2 August 1994/Accepted 19 September 1994

To maintain genomic integrity, double-strand breaks (DSBs) in chromosomal DNA must be repaired. In mammalian systems, the analysis of the repair of chromosomal DSBs has been limited by the inability to introduce well-defined DSBs in genomic DNA. In this study, we created specific DSBs in mouse chromosomes for the first time, using an expression system for a rare-cutting endonuclease, I-SceI. A genetic assay has been devised to monitor the repair of DSBs, whereby cleavage sites for I-SceI have been integrated into the mouse genome in two tandem neomycin phosphotransferase genes. We find that cleavage of the I-SceI sites is very efficient, with at least 12% of stably transfected cells having at least one cleavage event and, of these, more than 70% have undergone cleavage at both I-SceI sites. Cleavage of both sites in a fraction of clones deletes 3.8 kb of intervening chromosomal sequences. We find that the DSBs are repaired by both homologous and nonhomologous mechanisms. Nonhomologous repair events frequently result in small deletions after rejoining of the two DNA ends. Some of these appear to occur by simple blunt-ended ligation, whereas several others may occur through annealing of short regions of terminal homology. The DSBs are apparently recombinoenic, stimulating gene targeting of a homologous fragment by more than 2 orders of magnitude. Whereas gene-targeted clones are nearly undetectable without endonuclease expression, they represent approximately 10% of cells transfected with the I-SceI expression vector. Gene targeted clones are of two major types, those that occur by two-sided homologous recombination with the homologous fragment and those that occur by one-sided homologous recombination. Our results are expected to impact a number of areas in the study of mammalian genome dynamics, including the analysis of the repair of DSBs and homologous recombination and, potentially, molecular genetic analyses of mammalian genomes.

The repair of DNA double-strand breaks (DSBs) is important for the maintenance of genomic integrity in all organisms. In Saccharomyces cerevisiae, DSBs are repaired primarily by homologous recombination both in vegetatively growing cells and in cells undergoing meiosis (38). During meiosis, transient DSBs are induced at a number of positions known to be hot spots for recombination (5, 15, 53, 60). In vegetatively growing cells, mating-type switching, a DNA rearrangement that occurs through a gene conversion mechanism, is initiated by a DSB at the MAT locus (50). The DSB is introduced by the site-specific HO endonuclease (24). For mechanistic studies of DSB-promoted recombination, HO endonuclease expression systems have been frequently used to achieve cleavage at a number of other sites in both chromosomal and extrachromosomal DNAs (13, 23, 33, 34, 43, 46, 47, 52).

Two distinct homologous recombination pathways for the repair of DSBs have been proposed for S. cerevisiae, DSB repair and single-strand annealing (16). In the DSB repair pathway, a DNA substrate containing a DSB is repaired from an unbroken homologous DNA template (54). The unbroken substrate is the donor of genetic information, whereas the cleaved substrate is the recipient. By contrast, the single-strand annealing pathway has the requirement that both recombination substrates be cut at or near the regions of homology (13, 36, 52). The DNA ends provide an entry site for an exonuclease, so that single strands are exposed and available for annealing. In this pathway, recombination is nonconservative, such that both substrates can be considered recipients of genetic information.

In mammalian cells, analyses of the repair of DSBs have been limited by the inability to introduce defined DSBs in chromosomal DNA. However, studies using transfected DNA have demonstrated that both pathways of homologous recombination are operational (17, 19, 20, 22, 28–30, 51, 58). In contrast to S. cerevisiae, mammalian cells are very proficient at nonhomologous recombination. DSBs in transfected DNA are frequently repaired by end-joining mechanisms involving short sequence homologies at the DNA ends or simple ligation of the ends (44).

Gene targeting experiments provide the clearest demonstration of the preponderance of nonhomologous events in mammalian cells. Homologous integrations of transfected DNA into chromosomal DNA are often only a few percent or less of total integrations (6). Homologous integrations are stimulated by DSBs in transfected DNA, although they are still only a fraction of the total (17, 19, 58). Considering that the chromosome is the recipient of genetic information in gene targeting, it could be postulated that a DSB at the target chromosomal locus would stimulate gene targeting, utilizing either the DSB repair or single-strand annealing pathway of recombination.

To study the repair of DSBs introduced into DNA in vivo, we have recently expressed a rare-cutting endonuclease, I-SceI, in mammalian cells (45). I-SceI is a mitochondrial group I intron-encoded endonuclease from S. cerevisiae with an 18-bp, nonpalindromic recognition sequence (9). I-SceI promotes intron homing, a process which apparently proceeds by DSB repair (for a review, see reference 10). Plasmid recombination
using the single-strand annealing pathway has also been stimulated by DSBs introduced by I-SceI in yeast (41), mammalian (45), and plant (42) cells. There is no evidence that I-SceI endonuclease is involved in any step of recombination downstream of cleavage. Constitutive expression of I-SceI has been found to be nontoxic to mouse cells (45), presumably because there are no natural cleavage sites in the mouse genome or they occur infrequently. We have now used this expression system to introduce DSBs into I-SceI sites introduced into mouse chromosomal DNA. We have analyzed both homologous and nonhomologous repair events.

**MATERIALS AND METHODS**

Plasmid constructs and DNA manipulations. Plasmid pSlneo was constructed by inserting an I-SceI site into the unique NcoI site of pMC1neoA2 (56). The site was inserted by using the oligonucleotides 5’-CATGATCCCTGTATATGCTTA-3’ and 5’-CATGAGGATGACAGTTTACACACACA-3’, which were hybridized together prior to ligation. The oligonucleotides have the 18-bp I-SceI consensus sequence defined by Colleaux et al. (9) plus 5’CATG overhangs for ligation into the NcoI site. The insertion of the I-SceI site results in an altered C terminus of the Neo protein, beginning at amino acid 192.

For sequencing analysis, the neomycin phosphotransferase (neo) genes were amplified from 4 µg of genomic DNA, using the primers Neo1 (5’-GCCCATGATCCCTGCGGATCCATGACACCAACAA) and Neo2 (5’-AAACAGCTAAGACAGTTTACACACACA-3’), producing a 0.9-kb fragment. The amplified genes were cleaved to 0.7 kb with PstI-BamHI and expressed in *Escherichia coli* by subcloning into the PstI-BamHI sites of pSV2neoV22. The pSV2neoV22 plasmid is a derivative of pSV2neo (49) and carries the Tn5 bacterial promoter driving a neo gene (3). The neo gene in pSV2neoV22 is defective as a result of an insertion mutation at the NeoI site that can be distinguished from the mutation in the S1 neo gene. Bacterial colonies containing DNA amplified from G418’ cell lines were screened for ampicillin and kanamycin resistance. Amplified DNA derived from G418’ clones that contain two unarranged neo genes gave rise to both kanamycin-resistant and kanamycin-sensitive colonies in approximately equal numbers, whereas amplified DNA from clones that contain one unarranged neo gene (e.g., clone SE3) gave rise to kanamycin-resistant colonies. Amplified DNA derived from class 2 (see below) gave rise to kanamycin-sensitive colonies only, since the Neo2 primer binding site is lost in the rearranged gene.

Plasmid DNA was prepared from the transformants and subjected to DNA sequencing, using the primer SS1 (5’-TCGATACAGGGATGCTGCGGAAG) (31). Southern analysis was performed according to standard procedures (31). The I-SceI expression vector pCMV-I-SceI and the parental vector pCMV5 have been described previously (1, 45).

**Cell culture and transfections.** Mouse 3T3 cells were transfected by the calcium phosphate protocol as described previously (7) without reducing the percentage of CO2 in the incubator. Plasmid DNA is transcribed uncut except where noted. The S1E cell line was derived by cotransfection of 17 µg of pSlneo and 3 µg of pkghyg (55). Transformed cells were selected by using 100 µg of hygromycin B per ml starting 1 day posttransfection. Colonies were picked approximately 2 weeks posttransfection and subsequently expanded for Southern analysis.

The S1E cell line was subsequently transfected to generate G418’ colonies. Approximately 18 to 22 µg of DNA per plate was used for each transfection. The E clones were generated by transfection of S1E cells with 0.5 µg of the gel-purified neo 673-bp PstI-BamHI fragment, 2 µg of pUC19, and 17 µg of pCMV-I-SceI. The 2E and SE1 to SE6 clones were generated in experiment 1 (Table 1). Clones SE7 to SE17 were generated in experiment 2 (Table 1). The 4E clones were generated by transfection of the SpI-marked neo fragment (1 µg of the 685-bp PstI fragment) and 17 µg of pCMV-I-SceI. In the latter three experiments, pSV2his was included in the transfection (0.5 µg per plate). Duplicate plates of cells were used for G418 selection, and one plate was used for histidinol selection. In all cases, the plates selected by using histidinol became confluent at approximately the same time, indicating that none of the DNAs had a substantial inhibitory effect on transformation. Selection for G418’ was begun 24 h posttransfection, using 0.8 mg of G418 per ml. Clones were picked approximately 15 days after transfection. For the control transfections with pMC1neoA2, the total number of transformants was estimated by splitting the cells after transfection and, in addition in experiment 2, by transfecting 1/10 fewer cells and multiplying by 10.

**RESULTS**

**Design of a system to monitor the repair of DSBs.** To monitor the repair of chromosomal DSBs, we placed the 18-bp I-SceI site within the selectable neo gene (Fig. 1A). The site was inserted by ligating a 22-bp oligonucleotide with NcoI overhangs into the unique NcoI site in the neo coding region, disrupting the NcoI site and creating a frameshift. The mutated neo gene was then translated into the genome of 3T3 cells. The design of the experiment was to genetically detect cleavage of the I-SceI site and repair of the DSB by restoration of a functional neo gene through homologous recombination of the

<table>
<thead>
<tr>
<th>Transfected DNA</th>
<th>No. of G418’ colonies (% relative to pMC1neoA2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Expt 1</td>
</tr>
<tr>
<td>pCMV5</td>
<td>0, 0</td>
</tr>
<tr>
<td>pCMV-I-SceI</td>
<td>32, 35 (2.6)</td>
</tr>
<tr>
<td>neo fragment</td>
<td>ND*</td>
</tr>
<tr>
<td>neo fragment + pCMV5</td>
<td>0, 0</td>
</tr>
<tr>
<td>neo fragment + pCMV-I-SceI</td>
<td>172, 174 (13)</td>
</tr>
<tr>
<td>pUC18</td>
<td>0, 0</td>
</tr>
<tr>
<td>pMC1neoA2</td>
<td>-1,250 (100)</td>
</tr>
</tbody>
</table>

* SIE 3T3 cells were seeded at a density of 4 x 10³ cells per 10-cm-diameter plate 1 day prior to transfection. Duplicate plates of cells were transfected in each experiment for most samples, as indicated. Experiment 1, DNA per plate; pCMV5, 17 µg; pCMV-I-SceI, 17 µg; neo fragment + pCMV5, 1 µg of neo 685-bp PstI fragment, 2 µg of pUC19, 17 µg of pCMV5; neo fragment + pCMV-I-SceI, 1 µg of neo 685-bp PstI fragment, 2 µg of pUC19, 17 µg of pCMV-I-SceI; pUC18, 20 µg; pMC1neoA2, 3 µg of pMC1neoA2, 17 µg of pCMV5. Each plate, with the exception of the pUC18 plate, was also transfected with 2 µg of pSV2his as a transfection control. Experiment 2, DNA per plate; pCMV5, 17 µg of pCMV5, 3 µg of pUC19; pCMV-I-SceI, 17 µg of pCMV-I-SceI; 3 µg of pUC18; neo fragment, 1 µg of neo 685-bp PstI fragment, 20 µg of pUC18; neo fragment + pCMV5, 1 µg of neo 685-bp PstI fragment, 17 µg of pCMV5, 2.7 µg of pUC18; neo fragment + pCMV-I-SceI, 1 µg of neo 685-bp PstI fragment, 17 µg of pCMV-I-SceI, 2.7 µg of pUC18; pUC18, 20 µg; pMC1neoA2, 1 µg of pMC1neoA2, 2 µg of pUC18, 17 µg of pCMV-I-SceI. Each plate, with the exception of the pUC18 plate, was also transfected with 0.5 µg of pSV2his as a transfection control.

* ND, not done.
integrated neo gene with a transfected neo fragment. By introducing the homologous DNA as a linear fragment, recombination could occur by either the double-strand break repair or the single-strand annealing pathway.

The mutated neo gene was integrated into the genome of mouse 3T3 cells in a cotransfection with a hygromycin resistance gene. We identified several clones which carried the neo gene and expressed the 0.9-kb neo mRNA but were Neo−, as determined by sensitivity to the drug G418 (data not shown). Many clones had multiple copies of the 3.8-kb neo plasmid integrated. One clone, called S1E, was determined to have a simple integration pattern of the neo plasmid in its genome (Fig. 2). In this clone, two XhoI restriction fragments of 3.8 kb (neo gene A) and 2.5 kb (neo gene B) hybridize to the neo probe, indicating that two copies of the plasmid integrated in a head-to-tail fashion. One copy is the full-length 3.8-kb neo plasmid; the other contains a portion of the plasmid joined to chromosomal DNA. Each copy has a full-length neo gene, since the XhoI bands are reduced to 1.1 kb with the addition of BamHI. The structure is confirmed by an I-SceI digest, which produces a 3.8-kb band, and a double digest of XhoI-I-SceI, which produces a 0.9-kb band. As expected, NeoI does not cleave the XhoI fragments.

We focused our analysis on clone S1E because the two neo genes could be examined individually following transfection of the I-SceI expression vector. The G418 selection requires that one neo gene be repaired to produce a Neo+ phenotype. If the second neo gene is also cleaved at the I-SceI site, repair events at an unselected position can also be examined.

Transfections of the S1E cell line. To introduce DSBs at the I-SceI sites in the chromosomal neo genes, we used the expression vector pCMV-I-SceI (Fig. 1C). In pCMV-I-SceI, the I-SceI coding region has been fused to a simian virus 40 nuclear localization signal for efficient transport of I-SceI to the nucleus of mammalian cells (45). This fusion gene is expressed from the human cytomegalovirus promoter, as provided by the parental vector pCMV5 (1) (Fig. 1C). We have previously shown that pCMV-I-SceI functions well in 3T3 cells, resulting in efficient cleavage of transiently transfected DNA containing an I-SceI site (45).

The S1E cell line was transfected with pCMV-I-SceI or the parental vector pCMV5, in the presence or absence of a 0.7-kb 3' neo fragment that overlaps the cleavage site (Fig. 1A). No G418 colonies were obtained with pCMV5, the neo fragment, or the neo fragment with pCMV5 (Table 1). Apparently, gene targeting of this small fragment is inefficient. However, transfection of the neo fragment with pCMV-I-SceI resulted in a large number of G418 colonies. The 688 colonies in the two experiments that are presented in Table 1 represent approximately 12% the total number of transformants, as estimated from control transfections of pMC1neopA2. (The stable transformation efficiency of pMC1neopA2 provides only a relative measure in our experiments.) These results indicate that gene targeting is stimulated between 2 and 3 orders of magnitude by.
the introduction of a DSB at the neo locus in the S1E cell line. This result is not peculiar to this cell line, since we have obtained similar results with three other 3T3 cell lines containing mutated neo genes integrated into their genomes (data not shown). Surprisingly, transfection of pCMV-I-SceI alone also produced a significant number of G418® colonies. These represent 2.5% of the total number of transformants, or about one-fifth the number of colonies produced by the combined transfection of pCMV-I-SceI and the neo fragment. We have repeated these transfections several times with different preparations of pCMV-I-SceI. G418® colonies are always obtained in the absence of a neo fragment, although a cotransfected neo fragment always enhances the number of colonies.

Clones derived from transfection of pCMV-I-SceI alone. Seventeen G418® clones from the pCMV-I-SceI transfection were analyzed by Southern analysis. One of the two neo genes in each of the clones was found to have acquired the Neo® site and lost the I-SceI site (Fig. 3). The DNA sequence surrounding the I-SceI cleavage site was determined for the I-SceI-Neo® neo genes from five of the clones (5E2 to 5E6). Genomic DNA was amplified by using the primers Neo1 and Neo2 (Fig. 1A) and cloned into a vector which distinguishes in bacteria between amplified fragments which are Neo® or Neo® by their ability to confer kanamycin resistance. Approximately 60 bp of sequence surrounding the position of the original I-SceI site was determined, using the primer SS1 (Fig. 1A). The Neo® genes in each case were found to contain a wild-type neo sequence with an intact Neo® site. Considering that the insertion of the I-SceI site duplicated 4 bp of the neo coding region in the mutated neo gene, it appears that the DNA ends were rejoined at the 4-bp CATG repeats, regenerating the Neo® site and restoring neo function (Fig. 4).

Of the 17 clones analyzed from transfection of pCMV-I-SceI alone, one clone, 5E3, was found to have only one neo gene (Fig. 3 and data not shown). Clone 5E3 appears to have arisen from cleavage of both I-SceI sites, followed by joining of the 5' end of neo gene A with the 3' end of neo gene B. Cleavage of both sites results in a 3.8-kb deletion of the intervening chromosomal sequences.

The remaining 16 clones have two neo genes. In one of these clones, both neo genes are Neo® (5E12; Fig. 3B). In the other clones with two neo genes, neo gene A is Neo® in approximately half of the clones and neo gene B is Neo® in the other half, as distinguished by XhoI-Neo® digests (data not shown). The 15 clones with two different neo genes allow us to examine the fate of an I-SceI site at a position that did not undergo selection. The unselected neo gene in five of the clones, including clones 5E4 and 5E5 (Fig. 3), is I-SceI. The I-SceI sites in these genes either may have never been cleaved by I-SceI or were cleaved and then precisely religated. The other 10 clones, including clones 5E1, 5E2, and 5E6 (Fig. 3), have neo genes which are I-SceI. The unselected neo gene is not substantially altered in length in eight of the clones, including clones 5E2 and 5E6. By contrast, it is on a substantially different restriction fragment in two of the clones (5E1 and 5E17). For example, in clone 5E1, the XhoI fragment containing this altered gene is larger than in the parental cell line (data not shown).

The sequence surrounding the I-SceI cleavage site was determined for the I-SceI-Neo® Neo® genes from six of the eight clones with relatively minor lesions at the cleavage site (Fig. 4). Five of these were found to contain small deletions. Two clones have the four bases of the I-SceI overhangs deleted, whereas a third contains a deletion of the overhangs plus an adjacent base pair. The Neo® gene of another clone, clone 5E2, was found to contain a 9-bp deletion surrounding the cleavage site. Unlike the clones with the 4- and 5-bp deletions, the exact deletion endpoint in this clone cannot be determined, since it occurs within a 4-bp CCCT repeat. A fifth clone has a deletion of 29 bp surrounding the cleavage site. Similarly, there is an ambiguity in determining the exact
overhang (ATAA) of the bottom strand (Fig. 4). In another experiment, we have obtained a second clone with an insertion, and it is identical.

Clones derived from transfection of pCMV-I-SceI and the 3' neo fragment. Southern blot analysis was performed on 24 Neo⁺ clones originating from two independent transfections of pCMV-I-SceI and the 3' neo fragment. The probe specifically detects changes occurring at the target locus, since it is located 5' to the neo fragment (Fig. 1A). The clones can be grouped into two major classes based on the structure of the Neo⁺ gene. Southern analysis of representative clones from each class is presented along with the inferred genomic structure. Results are summarized in Table 2.

The Neo⁺ genes of the class 1 clones, like those of the 5E clones, are I-SceI− Neo⁺ and are contained on parental XhoI fragments of either 3.8 or 2.5 kb (Fig. 5). There are 11 clones in this class, comprising 46% of the Neo⁺ clones derived from the cotransfection. Considering that the clones derived from transfection of pCMV-I-SceI alone are 2.5% of the total number of transformants, and the class 1 clones are 5.5% of the total (12% × 0.46), the class 1 clones are likely generated both by homologous recombination with the neo fragment and by end joining, in approximately equal numbers. In subsequent experiments, we have marked the neo fragment with a restriction site polymorphism to specifically identify homologous recombinants (see below). The sequence surrounding the Ncol site was determined for eight of the class 1 Neo⁺ genes and was verified to be wild type in each case.

Of the 11 class 1 clones, two clones, E7 and E9, have only one neo gene, consistent with cleavage of both I-SceI sites prior to the repair event. Of the remaining nine clones, neo gene A is NcoI⁺ in six of the clones and neo gene B is NcoI⁺ in three of the clones. The unselected neo gene has lost the I-SceI site in five of the nine clones. In three of these clones, the XhoI fragment containing the I-SceI⁺ gene is similar in size to the parental S1E clone (e.g., clone 2E7; Fig. 5), indicating a minor lesion at the I-SceI site, whereas in the other two clones, the XhoI fragment appears substantially altered in size from the parental clone (e.g., clone 2E6; Fig. 5).

Using the approach described above, the sequence of the Neo⁻ gene was determined for clone 2E7. It was found to contain a 52-bp deletion surrounding the cleavage site (Fig. 4). As in two of the 5E Neo⁻ genes with deletions, there is an ambiguity in determining the exact deletion endpoint, since there is a 1 bp identity on both sides of the cleavage site.

The Neo⁺ genes of the second major class of clones are characterized as having acquired the Ncol site, although they are on XhoI fragments that are different in size from those of the parental S1E clone (Fig. 6). These class 2 clones have apparently undergone a homologous recombination on the 3' side of the NcoI site of the neo fragment and a nonhomologous event on the 3' side. The 5' event is necessarily homologous, as a result of the neo selection. Since the 3' end of the neo fragment extends beyond the neo stop codon, the 3' event is free for nonhomologous interactions, as long as the end of the neo gene is maintained. The 11 class 2 clones comprise 46% of the Neo⁺ clones derived from the cotransfection. No clones with this structure are derived from transfection of pCMV-I-SceI alone, indicating that their derivation depends on transfection of the homologous fragment. This is not unexpected, considering that end-joining events that do not utilize the CATG repeats would be unlikely to give a Neo⁺ phenotype.

One of the class 2 clones, clone 2E1, has only one neo gene. This clone may have arisen by cleavage of both I-SceI sites, followed by homologous recombination of the 5' end of neo gene A with the transfection fragment and nonhomologous

FIG. 3. (A) Southern analysis of genomic DNA from the 5E clones, which are derived from transfection of the S1E cell line with pCMV-I-SceI. Selected clones and digests are shown. (B) Deduced structures of the neo locus for 17 clones derived from transfection of pCMV-I-SceI. In this and subsequent figures, the Neo⁺ genes are shown with dark shading and the Neo⁻ genes are shown with light shading. For clones 5E1 to 5E6, the deduced structures are based on XhoI-BamHI-Ncol, XhoI-BamHI-I-SceI, XhoI-BamHI, and XhoI restriction analyses. For clones 5E7 to 5E17, the deduced structures are based on XhoI, XhoI-Ncol, and XhoI-I-SceI restriction analyses, to distinguish between neo genes A and B. The jagged 3' ends for the Neo⁻ genes in clones 5E1 and 5E17 indicate that the restriction fragments are substantially altered in size.

deletion endpoint in this clone. In this case, there is a 1-bp identity on both sides of the deletion endpoint.

By contrast, the Neo⁻ gene of clone 5E6 has an insertion of 1 bp. In this clone, it appears that an A was added to the 3'
FIG. 4. neo gene sequences around the I-SceI cleavage site. The 18-bp I-SceI site is shaded. Cleavage by I-SceI produces a 3' four-base overhang, as indicated. The deletion endpoints for Δ4 and Δ5 are clearly delineated, as indicated above the I-SceI site. The deletion endpoints for Δ1, Δ3, Δ9, Δ22, Δ29, and Δ52 are ambiguous, as indicated, as a result of repeats of either 1 bp (Δ1, Δ3, Δ29, Δ52) or 4 bp (Δ9, Δ22) on either side of the cleavage site. The 22-bp deletion gives rise to a Neo+ gene by loss of the inserted I-SceI site and one CAT repeat. The insertion of 1 bp at the I-SceI cleavage site is also indicated. Deletions: Δ4, clones 5E8 and 5E16, 4-bp deletions of the I-SceI overhangs; Δ5, clone 5E10, 5-bp deletion of the I-SceI overhangs plus an adjacent 1 bp; Δ1, clone 2E11, 1-bp deletion within the I-SceI overhangs; Δ3, clone E10, 3-bp deletion within the I-SceI overhangs; Δ9, clone SE2, 9-bp deletion at a CCC repeat; Δ29, clone 5E14, 29-bp deletion at a C repeat; Δ52, clone 2E7, 52 bp at an A repeat). The insertion of 1 bp occurs in clone 5E6.

recombination of the 3' end of neo gene B with the transfected fragment.

Of the remaining 10 clones that have two neo genes, six have the I-SceI site missing from the second gene. In four of these, the Neo+ gene is on a XhoI fragment that is substantially altered in size from the parental clone and was not further analyzed (data not shown). The DNA sequence was determined for two clones with relatively small lesions at the cleavage site. Clone E10 was found to have a 3-bp deletion at the I-SceI overhangs, and clone 2E11 was found to have a 1-bp deletion at the I-SceI overhangs (Fig. 4). In both cases, there is an ambiguity of 1 bp in determining the exact deletion endpoint.

Two other clones from these transfections were found to have Neo+ genes that were different in structure from the class 1 and class 2 clones. These genes are I-SceI- NcoI- and are contained on a parental-sized XhoI fragment (data not shown). Sequence analysis surrounding the cleavage site revealed that the Neo+ genes in these two clones have only one nucleotide change from the wild type, in both cases at the first position of the NcoI site. In one case, it is a C-to-T transition. In the other, it is a C-to-A transversion. The amino acid at this position is unaltered, since the nucleotide changes occur at a wobble position. We have not identified any more clones with this structure in subsequent transfections of a marked 3' neo fragment (see below), and thus we have been unable to determine if they were derived from end joining or from homologous recombination with the transfected fragment.

Clones derived from transfection of pCMV-I-SceI and a marked 3' neo fragment. To definitively establish that the Neo+ genes of class 1 clones can be generated by homologous recombination, the 3' neo fragment was marked with a restriction site polymorphism (Fig. 1B). An SspI site was introduced 22 bp downstream from the NcoI site by a T-to-C substitution. This is a silent mutation, since it occurs at the wobble position of an isoleucine codon.

Cotransfections of pCMV-I-SceI and the marked 3' neo fragment were performed in the S1E cell line. Southern analysis of four G418R clones is shown in Fig. 7. Two of the four clones, 4E205 and 4E208, are characterized as class 1 clones, since the Neo+ genes are I-SceI- Neo1+ and are on a parental-sized XhoI fragment (2.5 kb, for both clones).

<table>
<thead>
<tr>
<th>Transfected DNA</th>
<th>Total no. of clones</th>
<th>Neo+ gene</th>
<th>Neo+ gene</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Class 1</td>
<td>Class 2</td>
</tr>
<tr>
<td>pCMV-I-SceI</td>
<td>17</td>
<td>17*</td>
<td>0</td>
</tr>
<tr>
<td>neo fragment + pCMV-I-SceI</td>
<td>22d</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>SspI* neo fragment + pCMV-I-SceI</td>
<td>32</td>
<td>17*</td>
<td>15</td>
</tr>
<tr>
<td>SspI+</td>
<td>11</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>SspI+</td>
<td>7</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

* Class 1 Neo+ genes are NcoI+ and are on a parental-sized XhoI fragment. Class 2 Neo+ genes are NcoI+ and are on nonparental-sized XhoI fragments.

* Includes clones in which only one neo gene is present (e.g., E7) or in which both neo genes are apparently Neo+.

* Both neo genes in one clone contain the NcoI site and have class 1 structures.

* Two clones are not included since the Neo+ genes do not have the NcoI site and hence are not characterized as being class 1 or class 2.

* Both neo genes in one clone contain the NcoI site and have class 1 structures, one of which also contains the SspI polymorphism.
FIG. 5. Southern analysis of the class 1 clones derived from transfection of the S1E cell line with pCMV-I-Scel and the 3' neo fragment. Selected clones are shown, using XhoI (X), XhoI-NcoI (X/N), and XhoI-I-Scel (X/S) digestion. The deduced structures of the neo locus are presented, along with the names of additional clones which have the diagrammed structures.

Whereas the Neo+ gene of clone 4E205 is not cleaved by SspI, the Neo+ gene of clone 4E208 contains the SspI polymorphism, confirming that it was derived by homologous recombination with the marked 3' neo fragment. The other two clones, 4E206 and 4E207, have class 2 Neo+ genes, having acquired the NeoI site and lost the I-Scel site, yet residing on novel XhoI restriction fragments. Both contain the SspI polymorphism, as expected, since class 2 clones are not generated by transfection of pCMV-I-Scel alone (Table 2).

A total of 33 Neo+ genes were examined by Southern analysis (Table 2). Of the 15 that had class 2 structures, all were SspI+ as well as NcoI+. This finding confirms that all class 2 clones are generated by homologous recombination with the transfected fragment. Of the 18 with class 1 structures, 7 have acquired the SspI polymorphism. The 11 class 1 genes that are SspI- are generated either by end joining or by homologous recombination with a very limited gene conversion tract length. This issue can be addressed by reconstructing the mutated neo gene such that end joining cannot give rise to a G418R phenotype. In one of the clones, both neo genes have class 1 structures, one of which contains the SspI polymorphism.

Examination of the Neo+ genes in these clones further substantiates efficient cleavage of both chromosomal I-Scel sites. Three of the four clones shown in Fig. 7 have lost the I-Scel site from the second gene. Only clone 4E205 has retained the site. The Neo+ gene of clone 4E208 appears to have been generated by cleavage of both sites, with deletion of intervening sequences. Clones 4E206 and 4E207 both have the second neo gene on XhoI fragments that are different in size from the parental fragment. A summary of the results is presented in Table 2. In these experiments, the I-Scel site was present in the second neo gene in only 6 of 32 clones.

DISCUSSION

DSBs have been introduced into the genome of mouse 3T3 cells by using an expression system for a heterologous, rare-cutting endonuclease, I-Scel. A genetic assay has been devised to monitor DSBs, scoring for the restoration of a Neo+ phenotype as the result of repair of DSBs in either of two tandem neo genes mutated by insertion of an I-Scel cleavage site. Minimally, 12% of cells transfected with an I-Scel expression vector have at least one of the two I-Scel sites cleaved. Results from the 71 clones that we have examined suggest that cleavage of both sites occurred in more than 70% of selected clones. Cleavage of both sites in a portion of the clones deletes 3.8 kb of intervening chromosomal sequences.

Repair of the DSBs occurs both by end-joining mechanisms and by recombination with a transfected linear fragment. The end-joining clones are derived from transfection of the I-Scel expression vector and comprise 2.5% of stable transformants. The gene-targeted clones are derived from cotransfection of the I-Scel expression vector and a 0.7-kb homologous fragment and are approximately 10% of stable transformants. This represents a 2- to 3-order-of-magnitude stimulation of gene targeting, since gene targeted clones are rarely detected in the absence of endonuclease expression.

Our system selects one particular end-joining product that restores a functional neo gene through a 22-bp deletion. The presence of a second neo gene allows us to examine other end-joining products. Of 16 clones with two neo genes, 11 have lost the I-Scel site from the second neo gene. In 1 of the 11 clones, both neo genes are NcoI+, indicating that the selected product is a fraction of total possible end-joining products.
Most of the 10 Neo- genes have small lesions at the I-SceI cleavage site, and several of these are small deletions. Blunt-ended ligation may have given rise to the three Neo- genes that have deletions of the four-base I-SceI overhangs, with or without loss of an adjacent base pair. Analyses of other Neo- genes, as well as the Neo+ genes, suggest that another mechanism for the repair of chromosomal DSBs may involve short regions of terminal homology. For example, the Neo+ genes may be generated through the four bp CATG repeat resulting from insertion of the oligonucleotide. Upon cleavage by I-SceI, exonucleolytic digestion of both DNA ends may expose the four bases of the repeat, making them available for annealing and leading to a deletion of the inserted oligonucleotide (Fig. 8). Further processing and ligation steps would result in a wild-type neo gene. A similar end-joining mechanism through terminal homology can be postulated for repair of five of the Neo- genes, using either 4 bp of homology in one clone or 1 bp in the other four. Alternatively, exonucleolytic digestion of both strands may have occurred, followed by blunt-ended ligation. To distinguish end joining through short terminal homologies or blunt-ended ligation, the selection system can be modified such that the mutated neo gene does not contain the CATG repeat.

Extensive analysis of end-joining reactions of plasmid DNA have demonstrated that short sequence homologies at the termini play a role in the repair of extrachromosomal DSBs in vivo (44) and in DNA end joining as tested in cell extracts (35, 39). Our data support the relevance of these studies to the repair of chromosomal DSBs. Recent studies of rare nonhomologous repair events of chromosomal DSBs in S. cerevisiae have revealed joining through terminal sequence homologies (25), suggesting that a genetic analysis of the control of these events is possible. Short sequence homologies have also been implicated in coding joint formation in V(D)J recombination (reviewed in reference 27). Interestingly, terminal deoxynucleotidyltransferase expression appears to block homology-dependent joining events, even in the absence of N-region addition. The effect of terminal deoxynucleotidyltransferase expression on the repair of I-SceI generated chromosomal breaks can readily be tested.

Two other Neo- genes that have been sequenced have an insertion of an A at the 3' overhang of the cleavage site of the bottom strand. Some eukaryotic DNA polymerases have been shown to catalyze nontemplated nucleotide addition reactions, with a bias toward the use of dATP (8), suggesting a mecha-
nism for the insertion. Considering that the insertion duplicates the terminal nucleotide of the I-SceI overhang, it could have alternatively arisen from an annealing/fill-in mechanism that has been proposed for end joining of some yeast chromosomal breaks (25). In this case, the terminal AT nucleotides from the top strand would be removed, and then the terminal nucleotides of both strands would anneal, priming a fill-in synthesis.

Another mechanism for the repair of chromosomal DSBs is recombination with a homologous DNA fragment (gene targeting). The large stimulation of gene targeting that we see with the introduction of a DSB at the target neo locus implies that chromosomal DNA ends in mammalian cells are recombogenic. The efficiency of gene targeting in this system (approximately 10% of transfected cells) supports the notion that the homology search, even in such a complex genome as the mouse genome, is not rate limiting (61). We have detected two major classes of gene-targeted clones, those derived from two-sided homologous recombination and those derived from homologous recombination on one side and nonhomologous recombination on the other side. Interestingly, such one-sided events have previously been observed in other mammalian gene targeting systems in the absence of a DSB at the target locus (4, 11, 21, 57).

The enhanced gene targeting that we detect in this model system may have applications for creating targeted mutations at loci of interest. One round of gene targeting according to established procedures would be necessary to position an I-SceI site at the locus. Subsequent rounds may allow the placement of subtle mutations with relative ease. Such a two-step approach would be analogous to those developed for modification of the Caenorhabditis elegans and Drosophila genomes, where DSBs are introduced by transposable elements (14, 40). DSB-promoted gene targeting events in embryonic stem (ES) cells could further enhance the double-replacement protocol that has recently been developed by using both positive and negative selection (59). However, it remains to be determined if transient I-SceI expression will compromise the pluripotency of ES cells. Although constitutive I-SceI expression is not toxic to 3T3 cells (45), ES cells are expected to be more sensitive to genomic perturbations. In addition, a more direct measure of the utility of a DSB-promoted approach needs to be undertaken in ES cells, considering the high proportion of nonhomologous events that we detect through end joining and in the formation of class 2 clones. It is possible that the precision of DSB-promoted homologous events can be enhanced by using longer homologous fragments or circular substrates. In the long term, if sufficient precision of DSB-promoted gene targeting is achieved, more generalized reagents for the introduction of chromosomal DSBs could be extremely valuable for targeting loci in a single round.

The ability to introduce specific breaks into mammalian chromosomal DNA will be expected to impact a number of other areas of research in addition to homologous recombination and DNA repair. Endonuclease cleavage may assist chromosome fragmentation for the purpose of genome mapping or genetic analysis in vivo (2, 12, 18), as well as studies of chromosome stability (48) and cell cycle control following DNA damage (for example, see reference 26).

ACKNOWLEDGMENTS

We thank Peter Romanienko and Rich Bennett for providing assistance and materials, and we thank Dana Carroll, Jim Haber, and Lorraine Symington for thoughtful comments on the manuscript.

P.R. is a recipient of an EMBO long-term postdoctoral fellowship. M.J is a PEW Scholar for the Biomedical Sciences and recipient of the Frederick R. Adler Chair. This work was also supported by the NIH (grant P30-CA-08748-26) and the Beckman Foundation.

The first two authors contributed equally to this work.

REFERENCES


Thompson, S., A. R. Clarke, A. M. Pow, L. M. Hooper, and D. W. Melton. 1989. Germ line transmission and expression of a cor-