5012–5019 Nucleic Acids Research, 1995, Vol. 23, No. 24

Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells

Fatima Smih, Philippe Rouet, Peter J. Romanienko[1] and Maria Jasin*

Cell Biology and Genetics Program and [1]Molecular Biology Program, Sloan-Kettering Institute and Cornell University Graduate School of Medical Sciences, New York, NY 10021, USA

Received September 14, 1995; Accepted November 6, 1995

ABSTRACT

Double-strand breaks (DSBs) are recombinogenic lesions in chromosomal DNA in yeast, *Drosophila* and *Caenorhabditis elegans*. Recent studies in mammalian cells utilizing the I-*Sce*I endonuclease have demonstrated that in some immortalized cell lines DSBs in chromosomal DNA are also recombinogenic. We have now tested embryonic stem (ES) cells, a non-transformed mouse cell line frequently used in gene targeting studies. We find that a DSB introduced by I-*Sce*I stimulates gene targeting at a selectable *neo* locus at least 50-fold. The enhanced level of targeting is achieved by transient expression of the I-*Sce*I endonuclease. In 97% of targeted clones a single base pair polymorphism in the transfected homologous fragment was incorporated into the target locus. Analysis of the targeted locus demonstrated that most of the homologous recombination events were 'two-sided', in contrast to previous studies in 3T3 cells in which 'one-sided' homologous events predominated. Thus ES cells may be more faithful in incorporating homologous fragments into their genome than other cells in culture.

INTRODUCTION

The repair of DNA double-strand breaks (DSBs) is necessary for the maintenance of genomic integrity in all organisms. In *Saccharomyces cerevisiae*, repair of DSBs occurs primarily by homologous recombination and requires members of the Rad52 epistasis group (1). DSBs introduced by the mating type endonuclease HO at recognition sites located either at the endogenous *MAT* locus (2) or at heterologous loci have been found to stimulate homologous recombination (3). During meiosis transient DSBs occur at positions known to be hot spots for meiotic recombination and it has been suggested that DSBs may initiate all meiotic recombination events (4). DSBs in chromosomal DNA have also been found to be recombinogenic in other organisms. For example, transposon-induced DSBs in *Drosophila* (5) and *Caenorhabditis elegans* (6) are repaired by homologous recombination.

In mammalian cells, repair of DSBs involves both homologous recombination and non-homologous DNA end joining mechanisms. Evidence for both types of repair was first obtained in experiments with transfected DNA. DNA containing a variety of end configurations produced by restriction enzymes *in vitro* can be recircularized efficiently *in vivo* (7). Additionally, homologous DNAs will recombine efficiently when there are DSBs at or near the homology regions (8). Homologous recombination is stimulated by DSBs in transfected substrates whether the DSBs are introduced *in vitro* or *in vivo* (8,9).

A DSB in transfected DNA will also stimulate recombination between transfected DNA and chromosomal DNA (10–13). With a DSB in the transfected DNA the chromosome is formally the donor of genetic information to the recipient transfected DNA. Thus in gene targeting experiments the targeting vector becomes integrated upon marker selection. According to prevailing models, in which DSBs promote recombination, the chromosome would have to contain the DSB in order to become the recipient of genetic information (3,14).

Recently it has become possible to introduce DSBs into defined chromosomal loci in mammalian cells using the rare cutting endonuclease I-*Sce*I (15). The I-*Sce*I endonuclease cleaves an 18 bp non-palindromic site (16) and it has been found that constitutive expression of I-*Sce*I is non-toxic to mouse cells, presumably because there are no endogenous sites in the mouse genome or because they occur infrequently (9).

I-*Sce*I sites have been introduced into mammalian chromosomal DNA in mouse 3T3 cells, PCC7-S teratocarcinoma cells and Ltk$^-$ cells (15,17,18). Chromosomal DSBs introduced by I-*Sce*I at these sites were found to stimulate gene targeting in 3T3 cells and PCC7-S cells, but not in Ltk$^-$ cells. However, in 3T3 cells two thirds of the gene targeting events were found to be 'one-sided', such that one end of the transfected DNA integrated homologously, whereas the other end integrated non-homologously. We have now developed an ES cell expression system for I-*Sce*I and report on the repair of DSBs by gene targeting in these cells.

MATERIAL AND METHODS

Plasmid constructs and DNA manipulations

Plasmid pS2neo was generated by mung bean nuclease treatment of *Nco*I-cleaved pMC1neopA2 (19), followed by ligation of the

* To whom correspondence should be addressed

annealed 18 base oligonucleotides 5′-TAGGGATAA/CAGGGTAAT and 5′-ATTACCCTGTTAT/CCCTA. Integrity of the I-*Sce*I site insertion was checked by sequencing and *in vitro* cleavage with I-*Sce*I (Boehringer Mannheim). The V22*neo* gene was constructed by inserting the annealed oligonucleotides 5′-CATGCATTTCATTACCTCTTTCTCCGCACCCGACATAGATA and 5′-CATGTATCTATGTCGGGTGCGGAGAAAGAGGTAATGAAATG into the *Nco*I site of pMC1neopA2. This gene is not cleaved by I-*Sce*I.

The ES cell expression vector for I-*Sce*I, pgk-I-SceI, was constructed by deleting the *Nae*I–*Eco*RI human cytomegalovirus promoter fragment from pCMV-I-SceI (9) and inserting the mouse *pgk1* promoter fragment derived from pKJ-1 (20). The promoter was obtained by PCR amplification with oligonucleotides 5′-CGGAATTCCGATATCGATTTCTACCGGGTAGGGGAGGGCGCT and 5′-CGGAATTCCGTCGAAAGCCCGGAGATGAGGAAGAG.

Cell culture and transfections and PCR analysis

ES cell line E14 (21) was cultured in the presence of 10^3 U/ml leukemia inhibitory factor (22,23) (ESGRO, Gibco Life Science) on gelatin-coated plates. For transient transfection of ES cells 3×10^6 cells were plated on 10 cm dishes 24 h prior to transfection. Cells were transfected with 15 μg pgk-I-SceI and 15 μg either pS2neo or pV22neo by the calcium phosphate protocol as described previously (24) without reducing the percentage of CO_2 in the incubator. Plasmid DNA was harvested 48 h post-transfection for Southern blotting as described (9,25).

ES parental clone 12 containing a single copy of pS2neo was obtained by electroporating 20 μg pgkhyg (26) with 60 μg pS2neo into E14 ES cells as described (27). Hygromycin-resistant transformants were selected with 100 μg/ml hygromycin B starting 1 day post-transfection. For derivation of gene targeted subclones clone 12 was transfected using the calcium phosphate protocol. ES cells were seeded at a density of 4×10^5 cells/10 cm plate the day before transfection and were transfected with 17 μg pgk-I-SceI or pgklacZ, with or without 1 μg gel-purified *Pst*I 685 bp donor fragment from the 3′-end of the *neo* gene. For the pMC1neopA2 control 4.5 μg pMC1neopA2 was transfected along with a neutral pUC plasmid (pUCBM21) to bring the total amount of DNA to 20 μg. Plasmid pSV2his was used as a transfection control in preliminary experiments (0.5 μg/plate). Selection was started 24 h post-transfection with 200 μg/ml G418 (Geneticin; Gibco). Colonies were picked ~2 weeks later and further grown under selective conditions in 6-well dishes until harvest for genomic DNA analysis.

PCR products from the clones were generated with the primers Neo1 and Neo2 (9) and were digested with either 20 U *Sce*I, 20 U *Nco*I or 10 U *Ssp*I and loaded on a 1.2% agarose gel.

RESULTS

Design of a system to monitor DSBs in ES cells

We had previously mutated a *neo* gene for selection of I-*Sce*I-induced DSB repair events (15). The mutation consisted of an 18 bp insertion of the I-*Sce*I site flanked by a 4 bp duplication of *neo* gene sequences. Gene targeting events could be selected upon I-*Sce*I cleavage by restoration of a *neo*$^+$ phenotype. However, non-homologous end joining events at the 4 bp microhomology occurred at a low, but significant, level, complicating the analysis. Thus a new mutated *neo* gene was constructed specifically to select for I-*Sce*I-promoted gene targeting events.

As with the previously constructed S1*neo* gene, the new gene, called S2*neo*, has the 18 bp I-*Sce*I cleavage site inserted into the selectable *neo* gene of pMC1neopA2 at an *Nco*I site (Fig. 1A). However, prior to insertion of the I-*Sce*I site the 4 base overhangs of the *Nco*I site were removed with mung bean nuclease, resulting in both an insertion mutation (18 bp) and a deletion mutation (4 bp) (Fig. 1B). Since a stop codon is created in the *neo* coding sequence by the insertion, the gene is non-functional. With this design we expected that restoration of a functional *neo* gene following I-*Sce*I cleavage would be dependent upon recombination with a homologous *neo* gene fragment.

I-*Sce*I has previously been expressed in 3T3 cells from the human cytomegalovirus (hCMV) promoter in the vector pCMV-I-SceI (15). To express I-*Sce*I in ES cells it was necessary to replace the hCMV promoter with a promoter that expresses well in ES cells. We chose the mouse *pgk1* promoter (28). A PCR product containing the *pgk1* promoter from –521 to –18 relative to the translation start site was amplified from a pgkneo template (20). Since the *pgk1* gene has multiple transcription initiation sites located between –44 and –101 (20), it was expected that the –521 to –18 sequence would provide the necessary signals for transcription initiation. The amplified fragment was cloned into pCMV-I-SceI, resulting in plasmid pgk-I-SceI.

Expression of I-*Sce*I was checked by co-transfection of pgk-I-SceI with the S2*neo* gene. Transfection of pgk-I-SceI led to a similar or better level of cutting of the S2*neo* gene relative to transfection of pCMV-I-SceI in COS1 cells and a large increase of cutting in ES cells (Fig. 2 and data not shown). Another *neo* gene, V22*neo*, which does not contain the I-*Sce*I site, was uncut in these co-transfections.

Introduction of the S2*neo* gene into the genome of ES cells

The aim of our experiments was to determine if recombination of a transfected *neo* gene fragment would be stimulated by a DSB within the S2*neo* gene integrated into the genome of ES cells. The S2*neo* gene was introduced into ES cells through co-transfection with a hygromycin (*hyg*) resistance gene. PCR analysis was performed on hyg^R clones to determine which of the clones had also integrated the S2*neo* gene. Out of 18 hyg^R clones we identified seven that had at least one copy of the S2*neo* gene (data not shown). The other 11 clones apparently did not integrate the mutated *neo* gene with the *hyg* selection marker. The clones that gave a PCR product were further analyzed by Southern blotting. One of the clones, clone 12, was found to contain a single intact copy of the S2*neo* gene. The structure of the *neo* locus in clone 12 is shown in Figure 3.

Gene targeting in ES cells is stimulated by chromosomal DSBs

To determine if gene targeting would be stimulated by the introduction of DSBs into the *neo* locus we co-transfected the expression vector pgk-I-SceI with a *neo* gene fragment into clone 12. The fragment is 0.7 kb and is derived from the 3′-end of the *neo* gene (Fig. 1A). It is marked with a silent *Ssp*I restriction site polymorphism located 23 bp downstream from the *Nco*I site (Fig. 1B). Incorporation of the *Ssp*I polymorphism allows us to unequivocally assign gene targeting events and to begin to

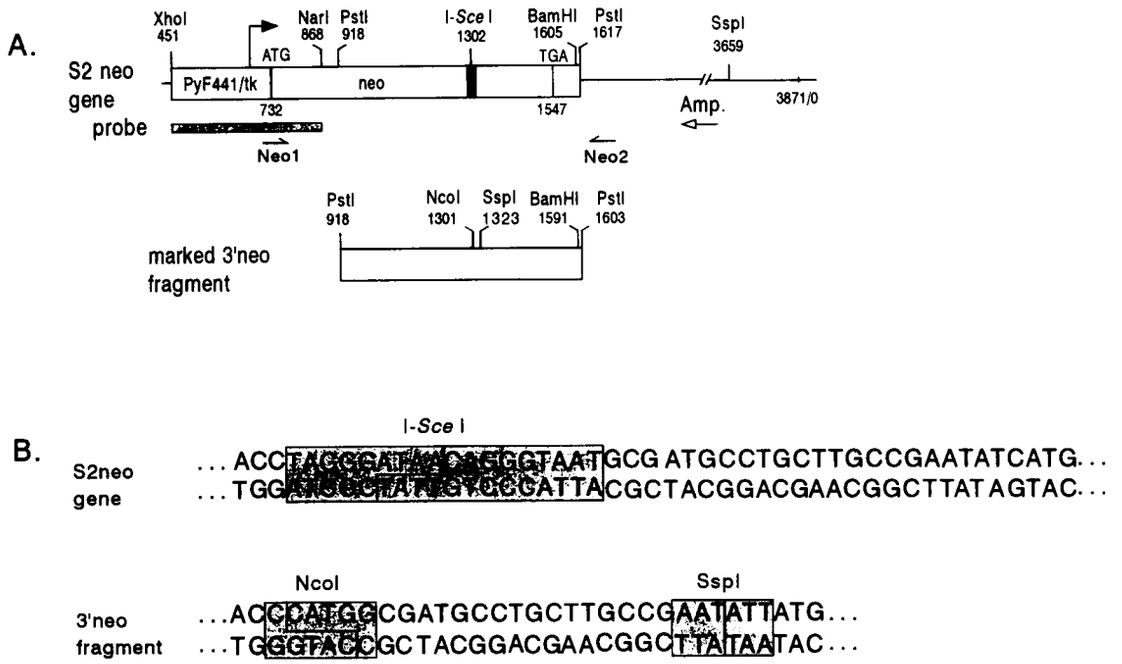

Figure 1. (**A**) Structure of the S2*neo* gene and the 3′ *neo* fragment. Insertion of the I-*Sce*I site into the *Nco*I site of the *neo* gene of pMC1neopA2 is indicated by the black bar. The promoter is derived from polyoma virus and the thymidine kinase gene (19). The position of relevant restriction sites and the primers used for PCR analysis (Neo1 and Neo2) are indicated. The probe for Southern analysis is the *Xho*I–*Nar*I fragment. The 3′ *neo* fragment has a wild-type *neo* gene sequence except that it is marked by a silent *Ssp*I restriction site polymorphism (15). (**B**) DNA sequences from the S2*neo* gene and the marked 3′ *neo* fragment. Cleavage by I-*Sce*I produces a 3′ four base overhang, as indicated. The first three bases of the I-*Sce*I site introduce a premature stop codon in the S2*neo* gene, rendering it non-functional.

Figure 2. *In vivo* cleavage of the S2*neo* gene at the I-*Sce*I site. The S2*neo* gene and the V22*neo* gene were separately co-transfected with pgk-I-SceI in ES cells. Two days post-transfection DNA was isolated from cells and restricted with *Xho*I and *Bam*HI *in vitro* and subjected to Southern analysis using a *neo* gene probe. Both S2*neo* and V22*neo* give 1.1 kb *Xho*I–*Bam*HI fragments. Cleavage of S2*neo* by I-*Sce*I expressed from pgk-I-SceI *in vivo* is detected by reduction of this fragment to 0.9 kb.

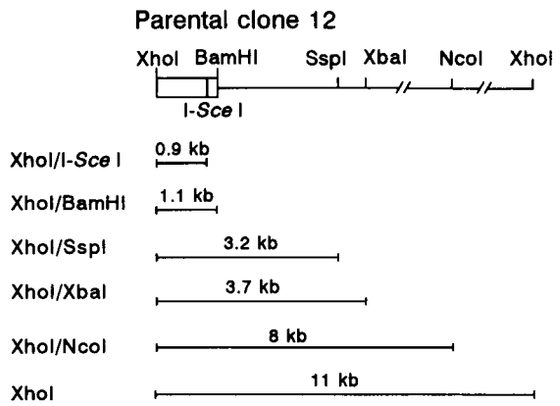

Figure 3. Structure of the S2*neo* locus in ES cell clone 12 as deduced by Southern analysis. Sizes of relevant restriction fragments are shown.

understand the mechanism of the recombinational repair of I-*Sce*I-generated DSBs.

Transfection of pgk-I-SceI alone into clone 12 did not generate any G418^R colonies (Table 1). Apparently, if cleavage occurs in the S2*neo* gene repair of the DNA ends in the absence of homology is unable to generate a functional *neo* gene. In contrast, when pgk-I SceI was co-transfected with the 3′ *neo* gene fragment a total of 50 G418^R colonies was obtained in two experiments. This number is 4.2% of the number of colonies obtained in transfection of pMC1neopA2 (1181 colonies), which gives an indication of the transfection efficiency in these experiments. No colonies were obtained in a co-transfection of the 3′ *neo* fragment in the absence of pgk-I-SceI, indicating that gene targeting of this short fragment is inefficient without a DSB at the target locus.

These gene targeted clones were generated by calcium phosphate transfection of ES clone 12. When an electroporation procedure was used G418^R colonies were obtained only rarely (data not shown). We interpret this result as being due to a much less efficient co-transfection of pgk-I-SceI with the 3′ *neo* fragment by electroporation than by calcium phosphate transfection. This point is currently being investigated.

Table 1. Gene targeting in ES cells

Transfected DNA	No. of G418R colonies[a]	
	Experiment 1	Experiment 2
pgk-I-SceI	0, 0	0, 0
neo fragment + pgk-I-SceI	7, 15	12, 16
neo fragment + pgklacZ	0, 0	0, 0
pMC1neopA2	230, 250	364, 336

[a]The two numbers for each experiment are the number of colonies obtained on two independently transfected plates of cells.

Analysis of the *neo* locus in the gene targeted clones

The gene targeted clones were subjected to PCR analysis of the *neo* locus. A 0.9 kb *neo* gene fragment was amplified from genomic DNA of 29 G418^R clones obtained from both experiments using the primers Neo1 and Neo2 (Fig. 1A). These primers are specific for the integrated S2*neo* gene and thus would not prime synthesis from a randomly integrated 3′ *neo* fragment. The amplified fragment was then digested with I-*Sce*I, *Nco*I and *Ssp*I. The fragment amplified from the parental clone is digested only by I-*Sce*I, as expected (Fig. 4, Pa). In contrast, none of the PCR products amplified from the 29 G418^R clones is cleaved by I-*Sce*I (Fig. 4 and data not shown). Rather, each of them is cleaved with *Nco*I, as would be expected from a gene targeting event with the transfected homologous fragment.

Cleavage with *Ssp*I allows us to examine incorporation of the silent restriction site polymorphism upon gene targeting. Of the 29 clones 28 have incorporated the polymorphism. One clone, clone 3, does not contain the 1 bp polymorphism, even though it has restored the *Nco*I site. Thus gene targeting frequently, but not always, involves gene conversion tract lengths at least as long as 23 bp.

PCR analysis examines a very defined region of the *neo* locus. To reveal any global modifications of the locus we performed Southern analysis on the G418^R clones using a probe derived from the 5′-end of the *neo* gene (Fig. 1A). Restriction digestion of genomic DNA from parental clone 12 with *Xho*I/*Nco*I gives an ~8 kb band (Fig. 5A). However, each of the G418^R clones has a 0.9 kb band with *Xho*I/*Nco*I digestion, indicating that the single *neo* gene in these clones has been converted to a wild-type sequence at the *Nco*I site (Fig. 5A and data not shown). Restriction digestion of the parental clone 12 with *Xho*I/*Ssp*I gives a 3.2 kb band (Fig. 5C). Each of the G418^R clones has a 0.9 kb *Xho*I/*Ssp*I band, with the exception of clone 3, which has the parental 3.2 kb band. These results confirm the PCR analysis and demonstrate that each of the G418^R clones has been gene targeted at the endogenous *neo* locus to contain the *Nco*I site and that all but one of the clones has also incorporated the *Ssp*I polymorphism.

To analyze further the structure of the targeted *neo* locus three other digests have been performed. Since the transfected homologous fragment is derived from the 3′-end of the *neo* gene, a homologous recombination event is only required at the 5′-end of the fragment to generate a functional *neo* gene. This has been termed a 'one-sided' homologous event. The 3′ event can be non-homologous, as long as the 3′-end of the *neo* gene remains intact. A *Bam*HI site is located downstream of the polyadenylation site for the gene, allowing us to examine the integrity of the event near the end of the transfected fragment. A *Xho*I/*Bam* HI digest gave the parental 1.1 kb band for each of the clones (Fig. 5E and data not shown). This suggests that no alterations occurred at the locus in the immediate vicinity of the transfected fragment.

To detect more global changes occurring at the locus *Xho*I/*Xba*I (Fig. 5F and data not shown) and *Xho*I/*Eco*RV digests were performed (data not shown). The parental clone gives a 3.7 kb band with *Xho*I/*Xba*I and a 5.5 kb band with *Xho*I/*Eco*RV. These bands are unchanged in 23 out of 29 clones. Thus the only apparent change in the *neo* locus in most of the clones is confined to the original position of the I-*Sce*I cleavage site.

A fraction of the clones, six out of 29 G418^R clones, have *Xho*I–*Xba*I fragments that are larger than the parental size. Clones 4 and 7 both have *Xho*I–*Xba*I fragments of ~6.5 kb, an increase of ~3 kb relative to the parental clone (Fig. 5F). Of the remaining four clones, two others have *Xho*I–*Xba*I fragments that are ~2–3 kb larger than in the parental clone, whereas the other two have increased more substantially in size, ~5 and 7 kb (data not shown). In all cases the *Xho*I–*Eco*RV fragments are similarly larger in size (data not shown).

Since transfected DNA can sometimes integrate into genomic DNA in large concatamers, Southern blots were reprobed with the 3′ *neo* fragment used to generate the targeted clones. If this fragment had integrated as a large concatamer the larger fragments in these six clones should give a much stronger hybridization signal than the parental sized fragments in the other clones. None of the six clones gave a stronger hybridization signal (data not shown), indicating that the increased size was not due to additional integration of the transfected fragment. Another possibility is that the size of these fragments is altered by integration of the I-*Sce*I expression vector. Integration of pgk-I-SceI is addressed in the next section.

Transient I-*Sce*I expression is sufficient to stimulate gene targeting

Since the G418 selection is started 24 h post-transfection, it is likely that transient I-*Sce*I expression is sufficient to stimulate gene targeting at the *neo* locus. To determine if pgk-I-SceI had integrated into the genome Southern blots were reprobed with I-*Sce*I coding sequences (Fig. 5B and D and data not shown). Only two of 26 clones analyzed had integrated the pgk-I-SceI plasmid. One had integrated a large number of copies (clone 17; Fig. 5B and D); the other had integrated one or two copies (data not shown). This result indicates that cleavage at the endogenous I-*Sce*I site occurs soon after transfection, during transient expression of I-*Sce*I from the expression vector, and that the expression vector need not be integrated to stimulate gene targeting.

DISCUSSION

We have designed a system to test whether chromosomal DSBs will stimulate gene targeting in ES cells. The cleavage site for the rare cutting I-*Sce*I endonuclease was introduced into the ES cell genome as an insertion mutation within a selectable *neo* gene. Co-transfections of a homologous *neo* gene fragment and an I-*Sce*I expression vector resulted in at least a 50-fold increase in targeted clones over transfections of the *neo* gene fragment alone, demonstrating that DSBs at a target locus in ES cells are

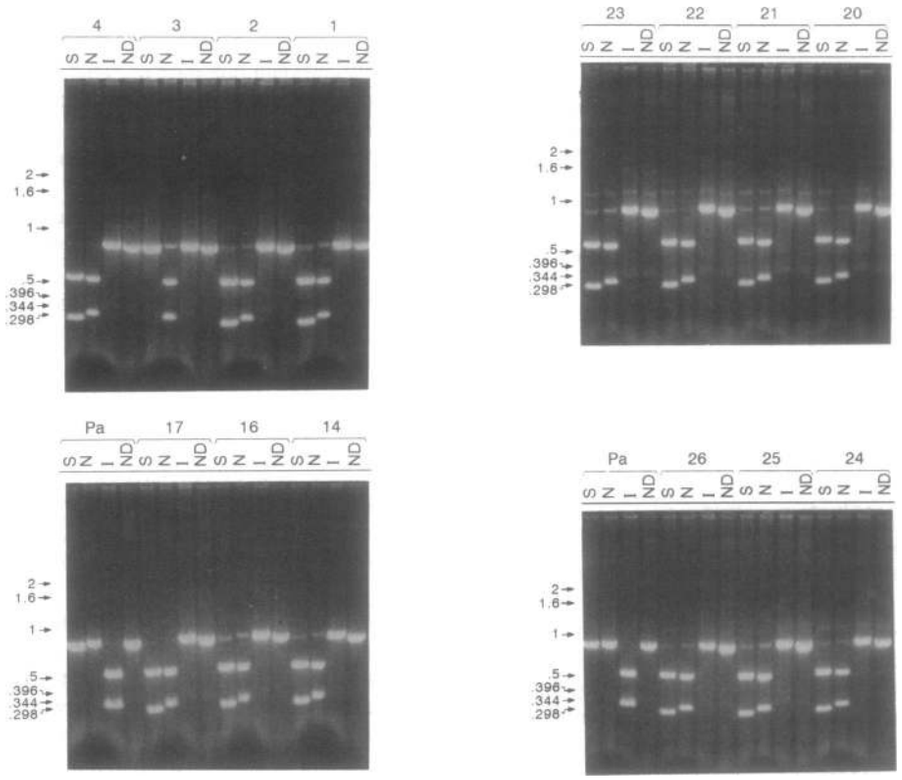

Figure 4. PCR analysis of parental clone 12 and G418^R subclones derived from co-transfection of pgk-I-SceI and the marked 3′ *neo* fragment. Fragments were amplified from genomic DNA using the primers Neo1 and Neo2 (Fig.1) and subjected to agarose gel electrophoresis with or without prior restriction digestion. The S2*neo* gene in parental clone 12 (Pa) is amplified as a 0.93 kb fragment . As expected, it is uncut by *Ssp*I and *Nco*I, but it is cleaved by I-*Sce*I to fragments of 0.59 and 0.34 kb. Conversely, the *neo* genes in G418^R subclones 1–4, 14, 16, 17 and 20–26 are uncut by I-*Sce*I but are all cleaved by *Nco*I. With one exception (sublone 3) they are also cleaved by *Ssp*I. S, *Ssp*I; N, *Nco*I; I, I-*Sce*I; ND, not digested.

recombinogenic. Transient I-*Sce*I expression is sufficient to stimulate gene targeting, since only a minority of clones have integrated the I-*Sce*I expression vector.

ES cells are now the fourth cell line to be tested for stimulation of gene targeting at chromosomal DSBs induced by I-*Sce*I. The other three are 3T3, Ltk$^-$ and PCC7 cells (15,17,18). In each of the cell lines tested, the I-*Sce*I site was first introduced into the genome within a defective selectable marker. Gene targeting was then assessed by transfecting a homologous fragment that could correct the I-*Sce*I cleavage site mutation. I-*Sce*I cleavage at the site has been brought about either through expressing I-*Sce*I *in vivo* or electroporating the purified enzyme. For three cell lines, 3T3, PCC7 and ES cells, a DSB at the target locus has been found to stimulate gene targeting 50-fold or more. Interestingly, DSB-promoted gene targeting events were not detected in Ltk$^-$ cells. Thus different cell lines may undergo recombinational repair of DSBs with different efficiencies.

In the current studies, most of the gene targeted ES cells were found to have undergone precise, 'two-sided' homologous recombination events. A fraction of the clones (21%) have undergone some other modification of the locus, leading to altered downstream restriction sites. As yet we have been unable to decipher precisely how these clones arose. They may be considered 'one-sided' homologous events, with homologous recombination occurring within the 5′-end of the *neo* fragment, reconstructing an intact *neo* gene, with some other event occurring at the 3′-end. These types of events have been seen in DSB-promoted gene targeting events in 3T3 cells (15), as well as in non-DSB-promoted targeting events (29–31). In 3T3 cells 'one-sided' events were found to be more frequent than 'two-sided' events, comprising as many as 68% of gene targeting events (15). These events were not screened for in PCC7 cells.

In each of the gene targeted ES clones the 18 bp I-*Sce*I site was removed and the 4 bp deletion of the *neo* gene was restored by recombination, resulting in a G418^R phenotype. A single base pair polymorphism located 23 bp from the I-*Sce*I cleavage site was incorporated in 28 of the 29 clones analyzed. This polymorphism results in a C→T transition. Incorporation of the polymorphism could be the result of heteroduplex formation between the chromosome and the transfected fragment, followed by mismatch correction (Fig. 6). Depending on which strands pair, the two mismatches possible are G–T and A–C. It has been shown in CV1 cells that >90% of G–T mismatches are corrected to G–C, whereas for A–C mismatches repair occurs equally in both directions (32). ES cells appear to have similar mismatch correction mechanisms (33). Based on these frequencies, we would have expected at least half of the clones to have retained the C/G base pair at this position. Since 97% of the clones have incorporated the polymorphism, either heteroduplex correction is occurring in the direction of the transfected fragment or there is no heteroduplex at this position. A non-randomness in mismatch correction has been postulated to occur at HO cleavage sites at the *MAT* locus in yeast and has been termed restoration (34). It has been postulated that the correction machinery senses the break in

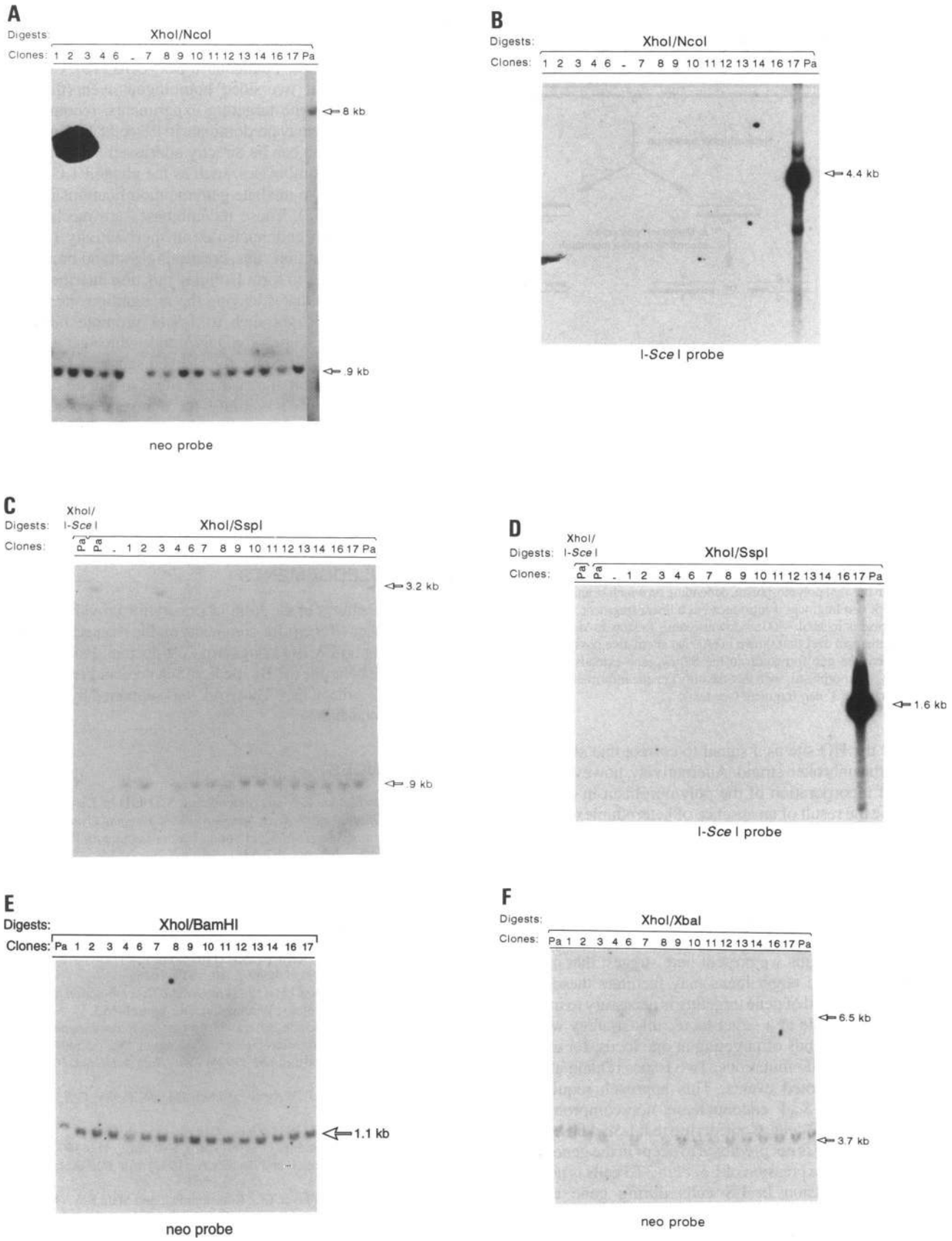

Figure 5. Southern analysis of digests of genomic DNA from parental clone 12 and G418^R subclones. Pa, parental clone 12. Numbers refer to the G418^R subclones. (**A**) *Xho*I–*Nco*I, *neo* probe. (**B**) *Xho*I–*Nco*I, I-*Sce*I probe. (**C**) *Xho*I–*Ssp*I, *neo* probe. (**D**) *Xho*I–*Ssp*I, I-*Sce*I probe. (**E**) *Xho*I–*Bam*HI, *neo* probe. (*F*) *Xho*I–*Xba*I, *neo* probe.

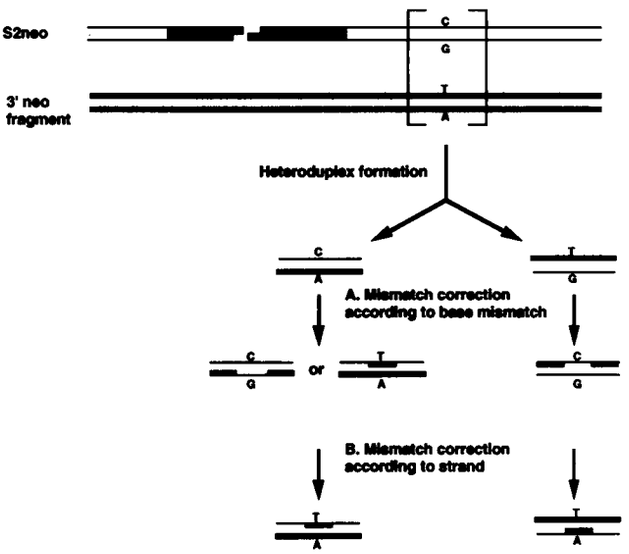

Figure 6. Heteroduplex formation at the *SspI* polymorphism. Both possible heteroduplexes are shown, as well as two types of correcting mechanisms. (**A**) Mismatch correction according to the type of base mismatch. In mammalian cells C–A mismatches are corrected with equal frequency to C/G or T/A, whereas T–G mismatches are predominately corrected to C/G. (**B**) Mismatch correction according to identity of DNA strand. Shown is a correction to the strand which is not broken (i.e. from the transfected DNA). The break is located 31 or 35 nt from the *SspI* polymorphism, depending on which strands are paired. (Note that the 3′ *neo* fragment is introduced as a linear fragment, although the break in this case is located ~300 nt downstream.) Results favor this type of correction mechanism over that shown in (A). An alternative possibility is that after I-*SceI* cleavage gap formation in the S2*neo* gene extends beyond the position of the polymorphism, such that the only genetic information available for repair is from the 3′ *neo* fragment (see text).

the strand at the HO site as a signal to correct that strand to the genotype of the unbroken strand. Alternatively, however, the high frequency of incorporation of the polymorphism in our experiments may be the result of an absence of heteroduplex formation. Gap formation may extend from the I-*SceI* cleavage site in the S2*neo* gene beyond the position of the polymorphism. Further analysis of the incorporation of polymorphisms upon DSB formation will allow us to distinguish these two mechanisms.

The pluripotential nature of ES cells has resulted in their use in numerous gene targeting experiments for the purpose of creating mutant mice. The results we present here suggest that introduction of DSBs into the target locus may facilitate these experiments. Since one round of gene targeting is necessary to introduce the I-*SceI* cleavage site at a target locus, this strategy would be suited to multiple rounds of targeting at one locus, for example, in the creation of subtle mutations. Two issues remain about the utility of DSB-promoted events. This approach requires that expression of the I-*SceI* endonuclease not compromise the pluripotency of the ES cells. Considering that I-*SceI* has an 18 bp site which statistically is not predicted to occur in the genome (16) and that constitutive expression of I-*SceI* in 3T3 cells is non-toxic (9), transient expression in ES cells during gene targeting experiments may be found to be non-deleterious. The second issue relates to non-homologous repair of DSBs. In addition to recombinational repair events, non-homologous end joining mechanisms are used to rejoin broken ends in mammalian cells (15,18,35,36). A comparison of the frequency of these two types of events has suggested that non-homologous events are somewhat more abundant than homologous events in 3T3 cells (15) and substantially more abundant in Ltk$^-$ cells (18). Considering the predominance of 'two-sided' homologous events that we have found in ES cell gene targeting experiments, recombinational repair mechanisms may predominate in ES cells. Using the I-*SceI* system this question can be directly addressed.

Site-specific recombinases, such as the phage P1 Cre protein, have been utilized to mediate genome modifications in ES cells and in mice (37,38). These recombinases are mechanistically quite distinct from endonucleases in their activity (39). They synapse two recognition sites, create single-strand breaks within the recognition sites to form Holliday junction intermediates and resolve the intermediates leaving the recognition sites intact. In contrast, endonucleases such as I-*SceI* promote homologous recombination by creating free DNA ends which act as initiators of recombination (40). Since the cleaved substrate is a recipient of genetic information, the cleavage site is lost upon recombination and there is no possibility for a reverse reaction. Thus expression of rare cutting endonucleases expands the repertoire of genome modifications that can be performed in mammalian cells. In addition to having potential as a tool in molecular genetic analyses, these endonucleases will also be useful for addressing basic biological questions about the nature of the repair of DSBs *in vivo*.

ACKNOWLEDGEMENTS

We thank members of the Jasin laboratory for providing assistance and J. P. Salier (Rouen) for comments on the manuscript. P.R. was a recipient of an EMBO Long-term Post-doctoral Fellowship. M.J. is a PEW Scholar for the Biomedical Sciences and recipient of the Frederick R. Adler Chair. This work was supported by the National Science Foundation.

REFERENCES

1 Petes,T.D., Malone,R.E. and Symington,L.S. (1991) In *The Molecular and Cellular Biology of the Yeast Saccharomyces: Genome Dynamics, Protein Synthesis and Energetics*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 407–521.
2 Strathern,J.N., Klar,A.J.S., Hicks,J.B., Abraham,J.A., Ivy,J.M., Nasmyth,K.A. and McGill,C. (1982) *Cell*, **31**, 183–192.
3 Haber,J.E. (1992) *Curr. Opin. Cell Biol.*, **4**, 401–412.
4 Wu,T.-C. and Lichten,M. (1994) *Science*, **263**, 515–518.
5 Gloor,G.B., Nassif,N.A., Johnson-Schlitz,D.M., Preston,C.R. and Engels,W.R. (1991) *Science*, **253**, 1110–1117.
6 Plasterk,R.H.A. (1991) *EMBO J.*, **10**, 1919–1925.
7 Roth,D.B. and Wilson,J.H. (1988) In *Genetic Recombination*. American Society for Microbiology, Washington, DC, pp 621–653.
8 Subramani,S. and Seaton,B.L. (1988) In *Genetic Recombination*, American Society for Microbiology, Washington, DC, pp 549–572.
9 Rouet,P., Smih,F. and Jasin,M. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 6064–6068.
10 Jasin,M., deVilliers,J., Weber,F. and Schaffner,W. (1985) *Cell*, **43**, 695–703.
11 Jasin,M. and Berg,P. (1988) *Genes Dev.*, **2**, 1353–1363.
12 Valancius,V. and Smithies,O. (1991) *Mol. Cell. Biol.*, **11**, 4389–4397.
13 Hasty,P., Rivera-Pérez,J. and Bradley,A. (1992) *Mol. Cell Biol.*, **12**, 2464–2474.
14 Szostak,J.W., Orr-Weaver,T.L., Rothstein,R.J. and Stahl,F.W. (1983) *Cell*, **33**, 25–35.
15 Rouet,P., Smih,F. and Jasin,M. (1994) *Mol. Cell. Biol.*, **14**, 8096–8106.
16 Colleaux,L., d'Auriol,L., Gailbert,F. and Dujon,B. (1988) *Proc. Natl. Acad. Sci. USA*, **85**, 6022–6026.
17 Choulika,A., Perrin,A., Dujon,B. and Nicolas,J.-F. (1995) *Mol. Cell. Biol.*, **15**, 1963–1973.

18 Lukacsovich,T., Yang,D. and Waldman,A.S. (1994) *Nucleic Acids Res.*, **22**, 5649–5657.
19 Thomas,K.R. and Capecchi,M.R. (1987) *Cell*, **51**, 503–512.
20 McBurney,M.W., Sutherland,L.C., Adra,C.N., Leclair,B., Rudnicki,M.A. and Jardine,K. (1991) *Nucleic Acids Res.*, **19**, 5755–5761.
21 Hooper,M., Hardy,K., Handyside,A., Hunter,S. and Monk,M. (1987) *Nature*, **326**, 292–295.
22 Smith,A.G., Heath,J.K., Donaldson,D.D., Wong,G.G., Moreau,J., Stahl,M. and Rogers,D. (1988) *Nature*, **336**, 688–690.
23 Williams,R.L., Hilton,D.J., Pease,S., Willson,T.A., Stewart,C.L., Gearing,D.P., Wagner,E.F., Metcalf,D., Nicola,N.A. and Gough,N.M. (1988) *Nature*, **336**, 684–687.
24 Chen,C. and Okayama,H. (1987) *Mol. Cell. Biol.*, **7**, 2745–2752.
25 Serghini,M.A., Ritzenthaler,C. and Pinck,L. (1989) *Nucleic Acids Res.*, **17**, 3604.
26 te Riele,H., Maandag,E.R., Clarke,A., Hooper,M. and Berns,A. (1990) *Nature*, **348**, 649–651.
27 Jasin,M. and Liang,F. (1991) *Nucleic Acids Res.*, **19**, 7171–7175.
28 Adra,C.M., Boer,P.H. and McBurney,M.W. (1987) *Gene*, **60**, 65–74.
29 Ellis,J. and Bernstein,A. (1989) *Mol. Cell. Biol.*, **9**, 1621–1627.
30 Berinstein,N., Pennel,N., Ottaway,C.A. and Shulman,M.J. (1992) *Mol. Cell. Biol.*, **12**, 360–367.
31 Jasin,M., Elledge,S.J., Davis,R.W. and Berg,P. (1990) *Genes Dev.*, **4**, 157–166.
32 Brown,T.C. and Jiricny,J. (1988) *Cell*, **54**, 705–711.
33 Steeg,C.M., Ellis,J. and Bernstein,A. (1990) *Proc. Natl. Acad. Sci. USA*, **87**, 4680–4684.
34 Haber,J.E., Ray,B.L., Kolb,J.M. and White,C.I. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 3363–3367.
35 Godwin,A.R., Bollag,R.J., Christie,D.-M. and Liskay,R.M. (1994) *Proc. Natl. Acad. Sci. USA*, **91**, 12554–12558.
36 Phillips,J.W. and Morgan,W.F. (1994) *Mol. Cell. Biol.*, **14**, 5794–5803.
37 Gu,H., Marth,J.D., Orban,P.C., Mossmann,H. and Rajewsky,K. (1994) *Science*, **265**, 103–106.
38 Smith,A.J.H., DeSousa,M.A., Kwabi-Addo,B., Heppell-Parton,A.,Impey,H. and Rabbitts,P. (1995) *Nature Genet.*, **9**, 376–385.
39 Stark,W.M., Boocock,M.R. and Sherratt,D.J. (1992) *Trends Genet.*, **12**, 432–438.
40 Dujon,B. (1989) *Gene*, **82**, 91–114.