Proc. Natl. Acad. Sci. USA
Vol. 85, pp. 8583–8587, November 1988
Genetics

Targeted mutation of the *Hprt* gene in mouse embryonic stem cells

(gene targeting/embryonic stem cells/homologous recombination/polymerase chain reaction amplification)

THOMAS DOETSCHMAN*, NOBUYO MAEDA†, AND OLIVER SMITHIES†‡

Laboratory of Medical Genetics and Genetics, University of Wisconsin, Madison, WI 53706

Contributed by Oliver Smithies, August 11, 1988

ABSTRACT **The hypoxanthine-guanine phosphoribosyltransferase (*Hprt*) gene has been mutated in mouse blastocyst-derived embryonic stem cells by site-directed homologous recombination. Embryonic stem cells were electroporated in the presence of a targeting DNA fragment containing two specific features: (*i*) The targeting DNA contained a promoterless neomycin phosphotransferase (*neo*) gene that, when located within the endogenous *Hprt* locus, could be transcribed from the promoter of the target locus. (*ii*) The targeting fragment had two short regions of homology with the endogenous *Hprt* gene: one, 132 base pairs long and the other, 1.2 kilobase pairs long. Targeted cells in which the designed homologous recombination event occurred were isolated either by selection with G418 followed by 6-thioguanine or by selection with 6-thioguanine alone. Even though <2 kilobases of homology existed between the exogenous and target DNAs, an average of 2.6 embryonic stem cells were successfully targeted for every 10^5 colonies surviving electroporation. Six of the *Hprt*$^-$ cell lines showed homologous recombination. These six lines were further analyzed by nucleotide sequencing a fragment that spans one crossover point after amplification by the polymerase chain reaction. Four lines had the expected sequence, whereas two lines had small deletions abutting the 132-base-pair region of homology.**

The development of gene targeting by means of homologous recombination between exogenous DNA and endogenous genomic sequences in somatic cells (1) and in mouse blastocyst-derived embryonic stem (ES) cells (2, 3) provides a powerful method for obtaining mutations specifically designed with respect to both type and location. Genetically modified ES cells can develop into complex embryonic structures in culture (4) and can produce transgenic animals after reintroduction into blastocyst-stage embryos (5, 6). Designed genetic modifications in ES cells *in vitro* should therefore be valuable for studying embryogenesis in cell culture and in the whole animal via transgenic mice.

We report gene-targeting experiments in which the hypoxanthine-guanine phosphoribosyltransferase (*Hprt*) gene was inactivated in ES cells by homologous recombination between the native *Hprt* gene and mouse exogenous DNA sequences containing the selectable *neo* (neomycin phosphotransferase) gene. The exogenous DNA contained short regions of homology with the target DNA as well as a promoterless *neo* gene. The promoterless *neo* gene was expected to substantially decrease the number of *neo*-expressing colonies formed by random insertion of the incoming DNA, while leaving unchanged the number of G418-resistant colonies due to homologous targeting. The fraction of G418-resistant colonies correctly modified should, therefore, increase considerably. To ensure that the hypoxanthine-guanine phosphoryltransferase (HPRT) enzyme encoded by the modified locus would be active, the exogenous DNA was arranged to provide a translational stop and reinitiation sequence at the beginning of the *neo* coding sequences. Six *neo*-positive, *Hprt*-negative ES cell lines were identified and analyzed. All of them resulted from homologous recombination; four were of the expected type, and two contained small deletions immediately upstream of the recombination site.

MATERIALS AND METHODS

Culturing, Electroporation, and Selection of ES Cells. ES-D3 cells (4) were cultured on embryonic fibroblast-feeder layers as described (4) with the following exceptions. The feeder layers used during selection were made from the mouse embryonic fibroblast cell line STO (gift from M. Evans, Cambridge, U.K.) (7); all feeder cells were irradiated [3000 rads (1 rad = 0.01 Gy) for embryonic fibroblasts or 4000 rads for STO cells] rather than treated with mitomycin C.

Cells were harvested with trypsin and electroporated as described (8) except that the field strength was 800 V/cm and the capacitance was 200 μF. The cells were at a concentration of 2×10^7/ml. Electroporation was done in the presence of 1.3 nM *Eco*RI-digested pTDA139 (see below) at room temperature in ES cell culture medium [high-glucose Dulbecco's modified Eagle's medium (GIBCO)/15% fetal bovine serum (Flow)/1×10^{-4} M 2-mercaptoethanol]. Two minutes after electroporation the cells were gently removed from the electroporation chamber and plated at 5×10^5 cells per 100-mm Falcon tissue culture dish preseeded with feeder cells. The feeder layers were either normal STO cells (already 6-thioguanine-resistant) or STO cells that we made G418-resistant by electroporation with pSV2neo (9). One day after electroporation the number of ES cell colonies growing on the feeder layers was counted to determine the colony-plating efficiency. At this time the ES cells on the G418-resistant STO cell feeder layers were exposed to medium containing G418 at 150 μg/ml for 5 days, and then the selection was changed to 6-thioguanine at 10 μM. Cells were not passaged at this time. Cells on normal STO cell-feeder layers were not treated with G418; four days after electroporation the ES cells were passaged. 6-Thioguanine selection was initiated one day later. The source plates for all passaged cells were noted to record the independence of selected colonies. After 11 days of 6-thioguanine selection, colonies were mechanically transferred with flame-drawn glass microcapillary pipettes into four-well (Nunc) trays containing buffalo rat liver cell-conditioned medium (10). Colonies that continued to grow for 2 more days under 6-thioguanine selection were then

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

Abbreviations: ES cells, embryonic stem cells; HPRT, hypoxanthine phosphoribosyltransferase; PCR, polymerase chain reaction.

*Present address: Department of Molecular Genetics, Biochemistry and Microbiology, University of Cincinnati College of Medicine, Cincinnati, OH 45267-0524.
†Present address: Department of Pathology and Program in Molecular Biology and Biotechnology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27514.
‡To whom reprint requests should be addressed.

passaged onto embryonic fibroblast-feeder layers for expansion and maintenance in the absence of selection.

Construction of Targeting Vector. The plasmid pTDA139 was the source of the targeting DNA and was constructed as follows. A 7-kilobase (kb) *Bam*HI fragment containing exons 2 and 3 of the mouse *Hprt* gene (see ref. 11 for mouse *Hprt* gene map) was isolated from a Charon 35 (12) genomic library of *Bam*HI-digested DNA from a female 129/Sv mouse and subcloned into pUC13 (13). A 2-kb *Bcl* I–*Bam*HI fragment containing a promoterless copy of the *neo* gene (the promoter region in pSV2neo is upstream of the *Bcl* I site) was prepared from pSV2neo and ligated via *Xho* I linkers into the unique *Xho* I site in the third exon of the 7-kb genomic *Bam*HI clone. The nucleotide sequence of the *neo* gene includes the tetranucleotide ATGA at positions 1–4 relative to the start of its translation. The length of the 5′ *Xho* I linker was chosen so that the second, third, and fourth of these nucleotides create a premature stop codon, TGA, in the *Hprt* reading frame. The first three nucleotides can, however, still serve as a normal methionine initiation codon, ATG, for the *neo* gene. We call this arrangement of the ATGA tetranucleotide a stop–setback–start construct. (A detailed analysis of *neo* expression from this type of stop–stepback–start construct will be presented elsewhere; see also ref. 14.) The sequence from the 5′ *Xho* I site to the *neo* start codon is given in Fig. 3.

Targeting (Exogenous) DNA. A 3.3-kb *Eco*RI fragment (see Fig. 1) excised from pTDA139 was used as targeting DNA. This fragment includes 132 base pairs (bp) of mouse genomic *Hprt* gene sequences immediately 5′ to the *Xho* I site in exon 3, the 2-kb promoterless *neo*-gene fragment from pSV2neo, and 1.2 kb of mouse genomic *Hprt* gene sequences immediately 3′ to the *Xho* I site.

Sequencing of Polymerase Chain Reaction (PCR)-Amplified DNA Fragments. PCR amplification was done as described (15). Twenty-nucleotide primers made on an Applied Biosystems DNA synthesizer at the Biotechnology Center, University of Wisconsin–Madison, were used without further purification except for one phenol and one chloroform extraction. The New England Biolabs PCR reaction mixture was modified to 67 mM Tris·HCl, pH 8.8/6.7 mM $MgCl_2$/16.6 mM $(NH_4)_2SO_4$/10 mM 2-mercaptoethanol/6.7 μM Na_2-EDTA/10% (vol/vol) dimethyl sulfoxide containing dATP/dCTP/dGTP/dTTP at 330 μM each, 1.5 μg of each primer, and 1 unit of *Taq* polymerase (*Thermas aquaticus* DNA polymerase; New England Biolabs) per 25 μl. From each targeted cell line, 2 μg of genomic DNA was amplified (4 × 500 ng in 25-μl reaction mixtures). Each sample was overlaid with 20 μl of light paraffin oil and amplified in an automatic temperature-shift apparatus for 40 cycles (one cycle = 15-sec melt time at 90°C and 5.5-min reaction time at 60°). An additional 1 unit of *Taq* polymerase was added after the first 20 cycles. The reaction products were pooled, ethanol precipitated, and then separated by 6% acrylamide gel electrophoresis. The amplified fragment was eluted from the gel by dialysis, ethanol precipitated, and washed. The fragments were end-labeled with [γ-^{32}P]ATP and T4 polynucleotide kinase and denatured on strand-separation gels; the individual strands were sequenced by the method of Maxam and Gilbert (16) with modifications as described (17).

RESULTS

Targeting Scheme. The overall scheme for the targeted mutation of the *Hprt* gene is diagrammed in Fig. 1. The fragment chosen for use as the transforming DNA (the 3.3-kb *Eco*RI fragment from pTDA139 as described) contains 2 kb of promoterless *neo*-gene sequences inserted into the *Xho* I site of the 1.3-kb *Eco*RI fragment from the *Hprt* gene that includes exon 3. This arrangement gives 132 bp of homology on the 5′ side and 1.2 kb of homology on the 3′ side of the

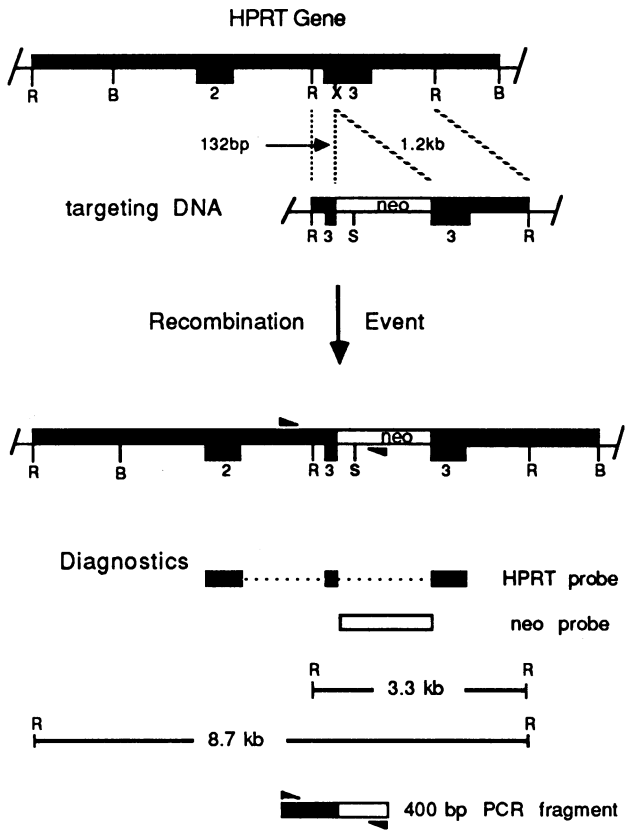

FIG. 1. Experimental design for homologous targeting of the mouse *Hprt* gene. The designed homologous recombination event requires the replacement of the endogenous 1.3-kb *Eco*RI fragment surrounding the third exon of the *Hprt* gene, with a 3.3-kb fragment from pTDA139, in which the 2-kb promoterless *neo* gene has been inserted into the *Xho* I site of the exon. Dashed lines define the lengths of homology between the target *Hprt* locus and the targeting DNA. Diagnostic restriction fragments and the 400-bp PCR-amplified fragment are shown to scale underneath the recombination scheme. For purposes of clarity the exon sizes are exaggerated. ■, HPRT-encoding DNA; ■, DNA exons; □, *neo* gene; R, B, and X, *Eco*RI, *Bam*HI, and *Xho* I sites; S, translational stop–stepback–start ATGA sequence; ▶, oligonucleotide primers for PCR.

inserted *neo* sequences. The targeting scheme is designed to replace exon 3 of the *Hprt* gene with the *neo*-disrupted exon 3 contained in the targeting sequences. The planned replacement event requires crossovers within the two regions of homology at the ends of the incoming DNA and will increase by 2-kb the length of restriction fragments spanning the recombination site. Genomic fragments and PCR-amplified fragments that can be used to diagnose the replacement event are shown (Fig. 1) and described in more detail below.

In our first experiment, 10^8 ES-D3 cells were electroporated in the presence of *Eco*RI-digested pTDA139 at 5 μg/ml, resulting in a molarity of the 3.3-kb fragment of 5 nM. After 24 hr the treated cells were selected with 400 μg of G418/ml for 16 days. Eleven G418-resistant colonies were obtained. About 2000 resistant colonies were obtained when the same number of cells were treated with pSV2neo under similar conditions. These numbers demonstrate that the promoterless *neo* gene construct reduces the number of G418-resistant transformants by >100-fold compared with that obtained with pSV2neo. This reduction supports our expectation that the promoterless gene will only be expressed at a restricted number of loci compared with pSV2neo. Nine of the 11 G418-resistant colonies were further selected with 6-thioguanine. Only one survived. Southern blot analysis (18) showed, however, that this G418/6-thioguanine-resistant cell was not targeted (data not shown), suggesting that the stop–

stepback–start *neo* gene in the *Hprt* locus might not be expressed sufficiently to protect the targeted cells from G418 at 400 μg/ml. Experiments were therefore carried out in which the ES cells were electroporated in the presence of the 3.3-kb targeting fragment and then either (*i*) doubly selected—first for 5 days with 150 μg/ml (a low level of G418 resistance that we found sufficient to kill ES cells after 14 days) and then with 6-thioguanine or (*ii*) selected with 6-thioguanine alone. Four 6-thioguanine-resistant colonies were obtained from the double-selection experiment, and five such colonies from selection with 6-thioguanine alone. Although both schemes yielded targeted colonies (see below), it is preferable to apply G418 selection for a short time before changing to 6-thioguanine selection for two reasons. First, some time is needed for residual, wild-type HPRT message and protein to be degraded by the targeted cells that can no longer synthesize the HPRT enzyme; when 6-thioguanine selection is applied too soon, residual HPRT can produce lethal levels of 6-thioGMP. Second, a short period of G418 selection slows ES cell growth sufficiently to eliminate the need for passaging cells before beginning 6-thioguanine selection.

Southern Blot Hybridizations Diagnostic for the Targeting Event. DNAs from the nine 6-thioguanine-resistant colonies were prepared and digested with *Eco*RI for analysis by Southern blot hybridization. Results after hybridization with an exon 3-specific probe are shown in Fig. 2*A*; Fig. 2*B* shows the results after hybridizing the same blot with a *neo*-specific probe. In lanes 4–6 and 11, a 3.3-kb fragment hybridized to both probes, but no 1.3-kb band was found—the expected pattern for the desired recombination event. The 6-thioguanine-resistant cells of lanes 2 and 9 are not HPRT-deficient from targeting for two reasons: (*i*) The exon-3 probe hybridizes only to the 1.3-kb fragment characteristic of the parental ES-D3 cells and the embryonic fibroblast-feeder cells (lanes 12 and 13, respectively); (*ii*) no hybridization is seen with the *neo*-specific probe. They are probably spontaneous HPRT mutants that had arisen during earlier passages of the cells. (We expected some spontaneous mutants because the starting ES cells had not been freed from HPRT-deficient cells by selection in hypoxanthine, aminopterin, and thymidine before the experiments.) One colony, represented in duplicate in lanes 7 and 8, is 6-thioguanine-resistant and also hybridizes to the *neo*-specific probe; yet, it has no *neo* insertion in exon 3 and requires further analysis.

The DNAs in lanes 3 and 10 contain an 8.7-kb fragment that hybridized to both probes. This 8.7-kb band suggests the occurrence of a homologous recombination event accompanied by loss of the *Eco*RI site at the 5′ end of the recombination zone (see Fig. 1). Hybridization of both probes to Southern blots of the same DNAs after digestion with *Hin*dIII, which does not cut within the recombination zone, detects a single 9-kb band in the same two samples, as well as in the other four targeted lines (data not shown). These data indicate that the two cell lines represented in lanes 3 and 10 were, indeed, modified by targeting. The loss of the *Eco*RI sites must, therefore, have been due to some additional event (see below). Thus, of the nine 6-thioguanine-resistant cell lines obtained, six were modified by targeting.

Characterization of the Loss of the *Eco*RI Site Adjacent to the Recombination Zone. To characterize the loss of the *Eco*RI site in the two lines represented by lanes 3 and 10 in Fig. 2, the region of DNA spanning the 5′ crossover sites in all six targeted ES cell lines was amplified by PCR, and the amplified fragments were then sequenced. The binding sites of the primers used and the location of the resulting amplified fragments are depicted at the bottom of Fig. 1. The primer sequences are shown in Fig. 3. Data show that the nucleotide sequences are as expected in the four cell lines with the *Eco*RI site but that the two clones lacking this site have small deletions (14 and 37 bp, respectively) immediately 5′ to the region of homology (Fig. 3).

Activity of the Promoterless *neo* Gene in the *Hprt* Locus. In the experiment with G418 preselection, all four 6-thioguanine-resistant lines subsequently obtained had been targeted. In the experiment without G418 preselection, two of the five 6-thioguanine-resistant lines had been targeted, but the other three lines appear to have resulted from spontaneous mutations. We questioned how much G418 resistance was provided by the *neo* gene incorporated into the *Hprt* locus in the stop–stepback–start configuration. The degree of G418 resistance conferred upon the targeted cells was therefore titered in buffalo rat liver cell-conditioned medium (which eliminates a small degree of G418 protection provided

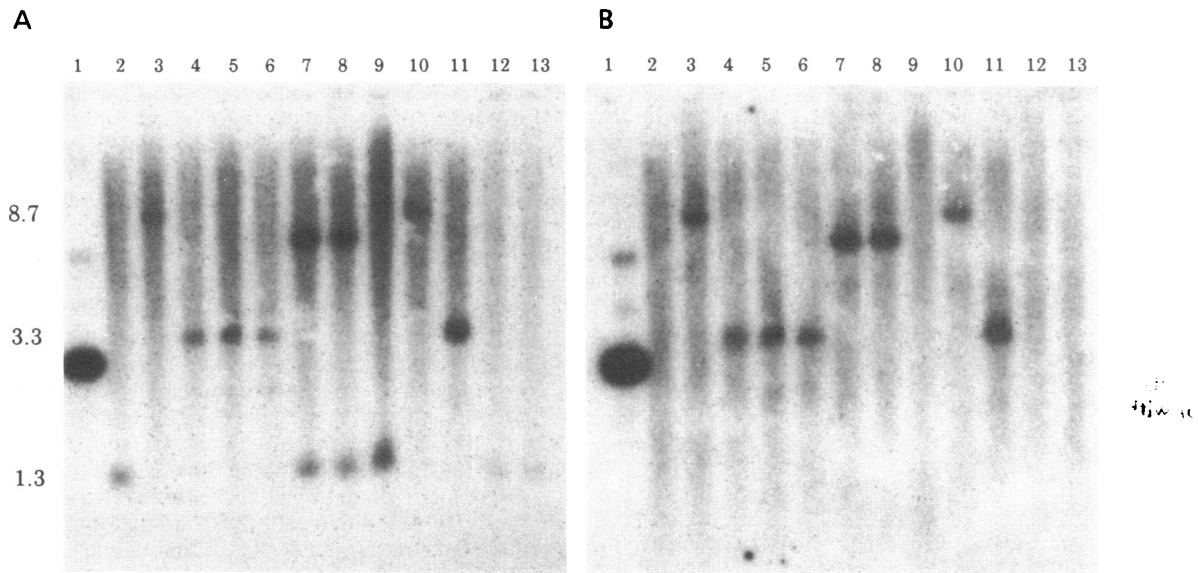

FIG. 2. Southern blot hybridizations of genomic DNA from 6-thioguanine-resistant colonies. Genomic DNA from 6-thioguanine-resistant and control cell lines was digested with *Eco*RI, and the blots were probed with a nick-translated human cDNA fragment-containing exons 2 and 3 of the *Hprt* gene (*A*), or pSV2neo (*B*). Lanes: 1, marker mix; 2–11, nine 6-thioguanine-resistant ES cell lines (lanes 7 and 8 contain DNA from the same cell line); 12, the parental cell line ES-D3; 13, primary embryonic fibroblast-feeder cells. Less DNA was loaded in lanes 12 and 13 than in the other lanes.

FIG. 3. Nucleotide sequences across the 5′ end of the recombination zone in six targeted cell lines. Two micrograms of genomic DNA from each of the six targeted cell lines were amplified by PCR to obtain the desired 400-bp fragment (see text). The 14- and 37-bp sequences that had been deleted in two targeted lines are underlined with bars. ~→, The two oligonucleotide primers used for PCR. Nucleotides 1–152 are sequences found exclusively in the endogenous *Hprt* locus and not in the targeting DNA. Nucleotides 153–400 were present in the targeting DNA. Sequences from the *Eco*RI to the *Xho* I sites (brackets over and under the sites, respectively) were common to both endogenous and targeting DNA. Sequences downstream of the *Xho* I site are from the *neo* construct. Uppercase letters designate coding sequences, and the upward arrowhead indicates the boundary between intron 2 and exon 3 of the *Hprt* gene. Translation of the *Hprt* gene (solid overline) runs through (broken overline) to a stop codon and then re-starts for the *neo* gene (solid underline).The stop–stepback–start tetranucleotide, ATGA, is boxed. The one-letter amino acid code appears either above or below the nucleotide sequence.

by *neo*-resistant feeder cells). The parental ES-D3 cells were found to be killed at a G418 concentration of 75–100 μg/ml. The cells modified by targeting died at a G418 concentration between 150 and 175 μg/ml. Thus, the *Hprt* promoter in conjunction with the stop–stepback–start arrangement of the promoterless *neo* gene provided an ≈2-fold increase in G418 resistance to the targeted cells.

Targeting Efficiency. There are several ways to express the frequency of successful targeting (Table 1). One way (corrected targeting efficiency, Table 1, column 5) is to note that for every 10^5 colonies that grew after electroporation and subsequent plating without selection, 1.3–5.3 colonies were from targeted cells. Another way to express the targeting frequency is with reference to the number of cells treated (overall targeting efficiency, Table 1, column 4). In this case the frequencies ranged from 0.4 to 1.6 targeted cells per 10^6 treated cells.

DISCUSSION

Two techniques useful for finding cells modified by gene targeting are the incorporation of an expressed drug-resistance gene as transfection tag and PCR screening of the transformed cells. Our results have implications for both techniques. We demonstrated a different way in which a drug-resistance transfection tag can be manipulated to limit irrelevant marker-positive transformations. Specifically, our construct contained a promoterless selectable marker gene that could be expressed from the promoter of the host target gene after targeting. Because most insertions of the promoterless construct at nonhomologous sites within the genome lack a promoter, the background of irrelevant drug-resistant transformants is greatly reduced. The level of protection against G418 derived from the *Hprt*-integrated *neo* gene is, however, less than that seen from a simian virus 40 promoter-driven *neo* gene. This difference may reflect the relatively modest transcription level of the *Hprt* gene, estimated to be expressed in mouse cells at levels such that ≈0.01% of mRNA molecules are *Hprt*-derived (19). Or the difference could also reflect a decrease in the stability of the fusion transcript because of its unusual translation pattern. In these experiments, the stop–stepback–start construction driven by the endogenous *Hprt* promoter could confer G418 resistance on ES cells of ≈50–75 μg/ml above the 75–100 μg/ml that they tolerate without the *neo* gene. By enriching for targeted recombinants among cells surviving G418 selection, this type of *neo*-gene construct may be useful for insertional mutagenesis of other expressed loci.

Table 1. Frequency of targeted mutation of the *Hprt* gene in ES-D3 cells

Treated cells	Surviving colonies*	Targeted cells	Overall targeting efficiency†	Corrected targeting efficiency‡
2.5×10^6	7.5×10^4	4§	1.6×10^{-6}	5.3×10^{-5}
5.0×10^6	15×10^4	2¶	0.4×10^{-6}	1.3×10^{-5}

*Colonies that formed after electroporation and replating. On the average, an estimated 3 colonies were formed from every 100 cells electroporated in the absence of DNA.
†Number of targeted cells per number of treated cells.
‡Number of targeted cells per number of colonies surviving electroporation and plating.
§Selection was first with G418 and then with 6-thioguanine.
¶Selection was with 6-thioguanine only.

Broad application of gene targeting depends on the development of techniques that allow cells genetically modified by gene targeting to be identified by screening rather than by selection. With current methods, ≈1 in 10^4–1 in 10^6 of the colonies surviving the entire transformation procedure become modified in the designated way (1–3). Detecting cells with the desired modification is consequently difficult unless the altered gene confers some selectable phenotype on the cell. PCR amplification is expected to help detect these rare events (15). A prerequisite for PCR amplification is that the oligonucleotide primers hybridize on the same DNA molecule sufficiently close to each other to allow amplification. Detection of a targeted recombinant sequence by PCR amplification requires that one of the two primers be chosen to hybridize with sequences specific to the target locus and the other be chosen to hybridize to sequences specific to the exogenous DNA. In this way only the recombinant fragment will be exponentially amplified. Currently, the practical limit of PCR amplification is ≈2 kb. Consequently, at least one homologous end of the integrating DNA must be short. This prerequisite was met by the 3.3-kb fragment from pTDA139, in which both homologous ends are shorter than 2 kb, one carrying only 132 bp and the other 1.2 kb of homology to the target locus. In experiments reported elsewhere (15), use of these *Hprt*-modified cells illustrate that the PCR method can detect a few targeted cells either alone or mixed with a 10,000-fold excess of nonrecombinant parental cells.

We estimated the frequency of gene targeting as the number of correctly modified, 6-thioguanine-resistant cells relative to the number of colonies that would have been obtained after electroporation and subsequent replating without selection. Under our experimental conditions, about half

of treated ES cells are killed by electroporation and about six colonies grow from every 100 cells that survive electroporation. We obtained similar numbers in earlier experiments involving the *Hprt* locus of ES cells (3). To compare these two data sets with each other and with the data of Thomas and Capecchi (2), who also targeted the *Hprt* locus, we applied the overall targeting efficiency as described and reported in Table 1. Overall frequencies of gene targeting are then 0.02–0.3 $\times 10^{-6}$ (2), 0.4–2.7 $\times 10^{-6}$ (3), and 0.4–1.6 $\times 10^{-6}$ in the present experiments. The amounts of homology to the target locus that were used were 3.7–9.7 kb (2), 3–5 kb (3), and in this work, 1.3 kb. Although increasing the amount of homology clearly increases targeting efficiency (2), it is uncertain why the 3–5 kb of homology in our earlier experiments (3) and the 1.3 kb of homology in these experiments targeted the *Hprt* locus at higher frequencies than those obtained by Thomas and Capecchi (2). Because the transformation efficiencies for pSV2neo and the rate of cell death caused by electroporation were roughly the same in both laboratories, technical factors, such as cell culture conditions or electroporation differences, probably did not play a significant role. Possibly the specific sequences of the exogenous DNA and/or the homologous part of the target locus will prove to be important.

The fidelity with which homologous gene-targeting events occur is of interest. Restriction mapping of modified target genes can assess fidelity, but such mapping is largely limited to detecting predicted fragments. At this level of detection, the homologous event usually appears to occur as planned. However, unless mutations affect gene function (20) or affect a restriction site diagnostic for the targeting event (as shown here), small insertions, deletions, or point mutations may go undetected. We detected a missing *Eco*RI site adjacent to the 5′-crossover point in two of the six targeted lines studied. The *Eco*RI restriction site at the equivalent point at the other end of the recombination locus remained intact during the recombination event. Sequencing of 400 bp spanning the 5′-crossover point showed four of the six targeted lines to be as expected, but two lines with missing *Eco*RI sites suffered small deletions. Sequences close to these deletions contained no repetitive elements to account for the deletions. A potential small stem-loop structure could have been formed within the sequences of the 37-bp deletion, but no such potential structure was detected within the sequences of the 14-bp deletion. Small deletions adjacent to partially homologous crossover points were reported by Anderson *et al.* (21), who analyzed six recombinant junctions formed when two cotransforming plasmids recombined with each other in mouse L cells. They found that the crossover points were in partially homologous areas in all cases (complete homology was not possible). In five of their six junctions, deletions were found immediately adjacent to the crossover points; in three of those five cases, a small insertion of 17–19 bp of DNA was sandwiched between the crossover point and the beginning of the deletion. Especially noteworthy in the present context was their observation that the lengths of the deletions came in multiples of ≈13 to 14 bp: 14, 27, 39 and 41, and 54. The two deletions that we found are 14 and 37 bp long, which suggests that they were generated by the same mechanism. Though neither Anderson and his collaborators nor we can yet describe the involved mechanism, their hypothesis is interesting—that a nucleic acid–protein complex partially unwinds both DNA helices to ≈13.5 bp per turn, thereby making the unwound helix vulnerable to nuclease activity at multiples of this helical repeat. Consistent with this hypothesis is data from an analysis of the dysmyelinating shiverer mouse mutation (22). The shiverer mutation contains a 13-bp deletion in one of the recombining strands between the breakpoint (which occurs at a region of partial homology) and the reunion site.

The presence of deletions immediately adjacent to the crossover were found only at the shorter end of homology in our experiments. This could be from chance or it may indicate that short homologous lengths are more error-prone. Several groups using plasmid × plasmid intermolecular (23) or plasmid intramolecular (24, 25) homologous recombination systems have reported that the efficiency of homologous recombination drops quite rapidly when the total length of homology between the recombining partners is <200–300 bp. Possibly, a protein–nucleic acid complex involved in the recombination requires a minimum length of homology to function properly. In our experiments total homology was 1.3 kb, and the efficiency of recombination was high, but possibly the shorter 132 bp of homology on one side induced occasional infidelities.

With respect to the precision of gene targeting, our data imply that restriction mapping alone is not always adequate to show whether the expected recombination event occurred. Nevertheless the data show that in many cases, the recombination event does occur exactly as experimentally designed.

We thank Natalie Borenstein and Debora Hautamaa for interpreting in blind tests the G418 resistance of the targeted cells, Susan McEvoy for help with the Southern blots, Mike Ganske for determining the transformation efficiency of ES cells using electroporation, Hyung-Suk Kim for assistance with the PCR amplification and nucleotide sequencing, Ed Shesely for valuable information with regard to infidelities in homologous recombination, and Tom Caskey (Houston) for human HPRT cDNA. This work was supported by National Institutes of Health Grants GM20069 and HL37001 to O.S. This report is paper 2995 from the Laboratory of Genetics, University of Wisconsin–Madison.

1. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. (1985) *Nature (London)* **317**, 230–234.
2. Thomas, K. R. & Capecchi, M. R. (1987) *Cell* **51**, 503–512.
3. Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S. & Smithies, O. (1987) *Nature (London)* **330**, 576–578.
4. Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W. & Kemler, R. (1985) *J. Embryol. Exp. Morphol.* **87**, 27–45.
5. Gossler, A., Doetschman, T., Korn, R., Serfling, E. & Kemler, R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9065–9069.
6. Robertson, E., Bradley, A., Kuehn, M. & Evans, M. (1986) *Nature (London)* **323**, 445–447.
7. Martin, G. R. & Evans, M. J. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1441–1445.
8. Boggs, S. S., Gregg, R. G., Borenstein, N. & Smithies, O. (1986) *Exp. Hematol.* **149**, 988–994.
9. Southern, P. J. & Berg, P. (1982) *J. Mol. Appl. Genet.* **1**, 327–341.
10. Smith, A. G. & Hooper, M. L. (1987) *Dev. Biol.* **121**, 1–9.
11. Melton, D. W., Konecki, D. S., Brennand, J. & Caskey, C. T. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 2147–2151.
12. Loenen, W. A. M. & Blattner, F. R. (1983) *Gene* **26**, 171–179.
13. Vieira, J. & Messing, J. (1982) *Gene* **19**, 259–268.
14. Peabody, D. S. & Berg, P. (1986) *Mol. Cell. Biol.* **6**, 2695–2703.
15. Kim, H.-S. & Smithies, O. (1988) *Nucleic Acids Res.*, in press.
16. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
17. Slightom, J. L., Blechl, A. E. & Smithies, O. (1980) *Cell* **21**, 627–638.
18. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
19. Melton, D. W., Konecki, D. S., Ledbetter, D. H., Hejtmancik, J. F. & Caskey, C. T. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 6977–6980.
20. Thomas, K. R. & Capecchi, M. R. (1986) *Nature (London)* **324**, 34–38.
21. Anderson, R. A., Kato, S. & Camerini-Otero, R. D. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 206–210.
22. Molineaux, S. M., Engh, H., De Ferra, F., Hudson, L. & Lazzarini, R. A. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 7542–7546.
23. Ayares, D., Chekuri, L., Song, K.-Y. & Kucherlapati, R. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 5199–5203.
24. Rubnitz, J. & Subramani, S. (1984) *Mol. Cell. Biol.* **4**, 2253–2258.
25. Chang, X.-B. & Wilson, J. H. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4959–4963.