MOLECULAR AND CELLULAR BIOLOGY, Jan. 1995, p. 45–51 Vol. 15, No. 1
0270-7306/95/$04.00+0

Analysis of Biological Selections for High-Efficiency Gene Targeting

KEITH D. HANSON AND JOHN M. SEDIVY*

Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Received 5 August 1994/Returned for modification 22 September 1994/Accepted 5 October 1994

A two-marker selection system that allows the efficient isolation of diploid gene knockouts by two sequential rounds of targeted homologous recombination has been developed. A systematic evaluation of the biological parameters that govern the selection process showed that a successful strategy must match the expression level of the target gene, the efficacy of the marker, and the selection stringency. An enrichment ratio of 5,000- to 10,000-fold, which resulted in a 30% targeting efficiency of the c-*myc* gene in a fibroblast cell line, has been achieved. Such efficiency brings the difficulty of gene targeting effectively down to the level of simple transfections, since only 10 to 20 drug-resistant clones need to be screened to recover several homologous hits. The general utility of the targeting strategy is of interest to investigators studying gene function in a large variety of mammalian tissue culture systems.

The principal impediment to the widespread application of targeted homologous recombination (gene targeting) for precise and specific genetic manipulations of mammalian cells is the low efficiency with which desired clones are recovered (30). In mitotic mammalian cells, nonhomologous recombination is 3 to 5 orders of magnitude more frequent than homologous recombination. Lack of success in significantly enhancing the frequency of homologous recombination has led to the development of genetic enrichments whose purpose is to eliminate clones arising from nonhomologous events.

The efficiency with which homologously targeted clones are recovered is the ratio of two absolute frequencies: that of homologous recombination at the target locus versus global nonhomologous recombination (30, 37). Locus-to-locus variability in the frequency of homologous recombination has been extensively documented. Evidence is accumulating that gene targeting in embryonic stem (ES) is significantly more efficient than in other cell types (1). The original gene-targeting strategies, either promoterless (11, 12, 21, 31) or positive-negative (PNS [23]), led to an explosion in research using ES cells. Analogous studies with other cell lines have remained largely inaccessible, mainly because the low intrinsic efficiency of most non-ES cell types makes the original gene-targeting protocols inadequate.

To illustrate the case, a typical gene-targeting experiment involves electroporation of vector DNA, selection of drug-resistant colonies, cloning and establishment of cell lines, and finally analysis by either Southern blotting or PCR to distinguish homologous and nonhomologous recombinants. In ES cells, targeting vectors in current use usually yield several homologous recombinants among the first 50 to 100 drug-resistant clones examined. Screening can be facilitated by combining individual clones into small pools for analysis by PCR. Screening above the 50- to 100-clone level requires a very significant expenditure of effort as well as time. Pooling large numbers of colonies followed by sib selection to recover clonal targeted cell lines is significantly hampered by clonal variability in growth rates. In cases where the targeted clones have reduced growth rates (32), recovery of clonal cell lines from complex pools is impossible. The targeting of the c-*raf* gene in our laboratory (28) is a typical example of the effort required with non-ES cell types. The primary screen involved ring cloning of 960 cell lines (40 24-well culture dishes) which were screened in pools of 12 by PCR (80 individual PCRs). The experiment resulted in the recovery of four targeted clones.

One strategy to circumvent the problem of low efficiency is to perform the targeting in ES cells and then establish cell lines of interest from the mutant animals (or embryos). There are, however, a number of compelling reasons why it would be advantageous to perform gene targeting directly in somatic cell lines. First, the generation of transgenic animals from ES cells is costly and requires specialized skills. Second, homozygous mutant embryos may not survive to a stage at which a cell line of interest can be established. Third, in many cases the establishment of cell lines is inefficient and difficult. Fourth, establishment of cell lines involves the poorly understood processes of crisis and immortalization. This means that any two cell lines, even if established from the same animal, cannot be considered isogenic. In contrast, a cell line and its directly targeted derivative are an isogenic pair. Fifth, many cell lines are already readily available and have been developed into extensively characterized model systems. Analysis of phenotypes elicited by gene targeting in such cell lines could thus draw on a large preexisting database. Sixth, the ES cell methodology is currently restricted to mice, a major impediment when human model systems are needed. For example, significant interspecies physiological differences call into question the validity of many extrapolations from studies of malignant transformation in rodent model systems to human carcinogenesis.

MATERIALS AND METHODS

Culture conditions. TGR-1 is an early-passage, 6-thioguanine (6-TG)-resistant subclone of the Rat-1 cell line (27). Cells were cultured in Dulbecco's modified Eagle medium (DMEM; Gibco) with glutamine, pyruvate, and high glucose, containing 3.7 g of sodium bicarbonate per liter, penicillin-streptomycin, and 10% calf serum, in a 5% CO_2 atmosphere at 37°C.

Selection conditions. All selections were performed by incubating electroporated cells for 48 h in the absence of selection, followed by trypsinization and plating directly into selective medium. Cells were plated in selective medium at

* Corresponding author. Mailing address: Department of Molecular Biophysics and Biochemistry, Yale University, P.O. Box 208024, 333 Cedar St., New Haven, CT 06520-8024. Phone: (203) 785-5429. Fax: (203) 785-6404. Electronic mail address: Sedivy@Biomed.med.yale.edu.

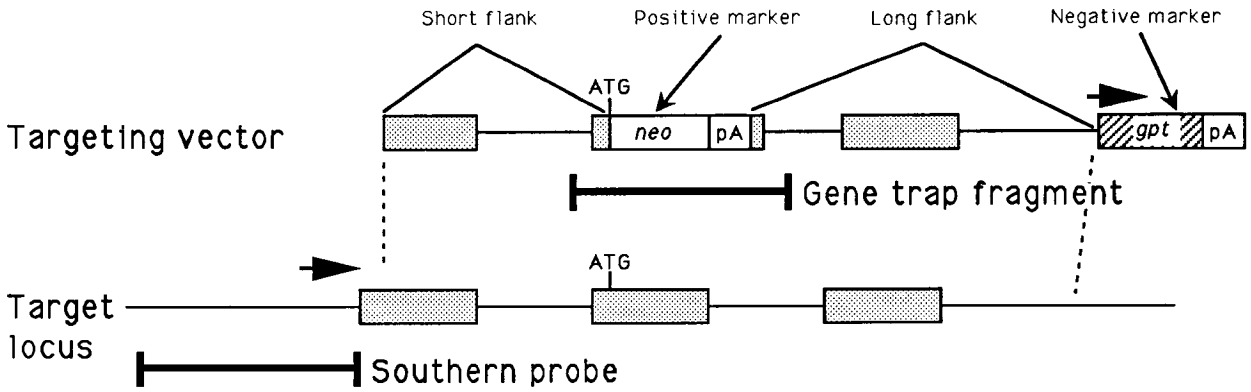

FIG. 1. Structure of the c-*myc* gene and design of targeting vectors. The c-*myc* gene contains three exons, and the Myc protein is initiated in exon 2. A *neo* promoterless construct is shown in the diagram (for details, see Materials and Methods). The *hyg* and *his* genes were substituted at the same position. In PNS vectors, the positive marker was the *hyg* gene driven by the Moloney murine leukemia virus long terminal repeat and was also substituted into exon 2. pA, polyadenylation site; heavy arrow, promoter; dashed lines, crossovers.

various dilutions but never exceeding 30% confluence at the time of plating. G418 (Geneticin) was from Gibco, hygromycin B was from Calbiochem, and histidinol · 2HCl and 6-TG were from Sigma. Gancyclovir was kindly provided by Syntex Laboratories, Palo Alto, Calif. Various drug concentrations were used, as indicated. Drug concentrations are expressed as total weight of the drug preparation. According to the data sheets of the manufacturer (Gibco), the fraction of active G418 varied from 45 to 55% in the lots used during this investigation. The His selection was performed by using histidine-free DMEM (Gibco custom preparation).

Electroporation conditions. Electroporation conditions have been described elsewhere (27). The amount of targeting vector DNA was held constant in all experiments (total of 10 μg in a 0.5-ml electroporation reaction).

Determination of enrichments. Enrichments of PNS constructs were determined by placing aliquots of a single electroporation under two selections: one for the positive marker only, and the other for both the positive and negative markers. A wide range of dilutions was plated in all cases. In the case of the diphtheria toxin gene (*dta*), parallel electroporations were performed by using identical constructs with and without *dta*. After 14 days at 37°C (feeding every 4 days), dishes were stained with crystal violet and colonies were counted.

Selection and analysis of recombinants. Individual drug-resistant colonies were picked by using cloning rings (Corning) and trypsin, transferred into 24-well culture dishes, and expanded into cell lines. In all targeting experiments, clonal cell lines were individually analyzed by Southern blotting hybridization. Genomic DNA was prepared from one 6-cm-diameter dish by a miniprep procedure involving sodium dodecyl sulfate lysis, proteinase K digestion, phenol extraction, and ethanol precipitation. Southern blotting analysis was performed by standard procedures (29), using a 5′ flanking c-*myc* chromosomal probe (Fig. 1 and 2). Since the probe sequences are absent from the targeting vector, nonhomologous recombination events are not detected. Neomycin phosphotransferase gene (*neo*) and histinol dehydrogenase gene (*his*) probes were also used, as indicated. Quantitative Western blotting (immunoblotting) analysis was performed as described previously (17, 27). Anti-Neo antibody was from 5′-3′, Inc. Total cytoplasmic extracts were prepared by the method of Nonidet P-40 lysis (27), and protein content was determined using the micro bicinchoninic acid assay (Pierce). Signals were detected by the enhanced chemiluminescence method (Amersham). Signals were quantitated by reference to lanes loaded with a dilution series of known quantities of purified Neo protein (5′-3′, Inc.).

Selection and analysis of mitotic nondisjunction events. TGR-1 cells were either infected with the pZIP*gpt* retrovirus vector (9) or electroporated with pSV2*gpt* plasmid DNA (25). Clones were selected in hypoxanthine-aminopterin-thymidine (HAT) medium (Sigma), and cell lines that contained single-copy inserts (determined by Southern blotting hybridization) were selected for 6-TG resistance. To determine the reversion frequency, the cells were grown for two passages in HAT medium and one passage (four to five doublings) in HT medium (HAT without aminopterin) and finally selected with 6-TG (6 μg/ml) in DMEM. For each cell line, a total of 2×10^7 cells were selected in 20 10-cm-diameter dishes (less than 30% confluent at the time of plating). After 14 days at 37°C (with feeding every 4 days), dishes were stained with crystal violet and colonies were counted. The frequencies reported were calculated by dividing the number of colonies by the number of cells plated.

Targeting constructs. All targeting constructs were of the replacement (double-crossover) type, were made with isogenic DNA, and contained identical c-*myc*-homologous regions (2 and 5.2 kbp on the 5′ and 3′ arms, respectively). The 5′ flank extended to the start of exon 1 (both P1 and P2 promoter sequences were absent), and the 3′ flank extended some 3 kbp past the end of exon 3. In promoterless vectors, the *neo* (34), *hyg* (hygromycin phosphotransferase [3]), or *his* (18) coding sequence was substituted directly at the ATG start codon of c-*myc*. The constructs thus encode native Neo, Hyg, or His protein. In all cases, the junctions were synthetic double-stranded oligonucleotides, which were ligated with appropriate restriction fragments from a genomic c-*myc* clone. PCR was not used in the constructions, and all relevant junctions were sequenced to confirm their fidelity. Negatively selectable genes *gpt* (guanine phosphoribosyltransferase [25]), *tk* (herpes simplex virus thymidine kinase [23]), and *dta* (38) were placed in various combinations on the flanks. In PNS vectors, the positive marker was the *hyg* gene driven by the Moloney murine leukemia virus long terminal repeat and was also substituted into exon 2 of c-*myc*.

RESULTS

The PNS selection produces only limited enrichments. The key feature of the PNS strategy is a negatively selectable marker on the flank of the targeting vector which is recombined away during the process of homologous integration. The enrichment of a PNS selection is the ratio of clones recovered with the positive selection only (total number of integration events) versus the combined positive and negative selections (homologous integrations plus all other events accompanied by inactivation of the negative marker). In a targeting experiment using a PNS construct, we tested 187 clonal cell lines by Southern blotting, but no targeted events were recovered (Fig. 1; Table 1; see Materials and Methods for experimental details). The degree of enrichment provided by the PNS selection was subsequently tested with numerous constructs, using positive (*neo*, *hyg*, and *his*) and negative (*gpt*, *tk*, and *dta*) selectable genes in various combinations. We never observed an enrichment exceeding 8- to 10-fold, and in most cases the enrichments were 2- to 3-fold. We did not see any significant differences among the various negative markers or with various combinations of positive and negative markers. When two distinct negative markers were used in one vector, the effects were roughly additive. The low enrichments can be explained by nucleolytic damage to the negative marker during electroporation (16). A large number of reports confirming these results have appeared in the literature (37).

Cells targeted with a promoterless *neo* vector are highly G418 resistant. The promoterless strategy uses vectors in which the positively selectable marker lacks its own promoter but its expression is activated from the target gene promoter following homologous integration. We succeeded in targeting c-*myc* with a promoterless *neo* vector; 72 cell lines were screened, and three hits were recovered, an efficiency of 4% (Fig. 2A; Table 1). The fidelity of recombination events was verified with several restriction enzymes and with *neo* as well as

TABLE 1. Summary of c-*myc*-targeting experiments[a]

Expt	Type of selection	Selectable genes	Selection (μg/ml)				No. of hits/total no. of clones analyzed	Targeting efficiency (%)
			G418	Hyq	6-TG	His		
1	PNS	*hyg gpt*		110	6		0/187	<0.5
2	Promoterless/PNS	*neo gpt*	600		6		3/72	4
3	Promoterless/PNS	*neo*gpt*	600		6		16/173	9
4	Promoterless/PNS	*neo*gpt*	3,000		6		6/19	32
5	Promoterless/PNS	*his gpt*	600		6	43	7/23	30
6	Promoterless/PNS	*hyg gpt*	600	110	6		0/48	<2

[a] Selection conditions are given in Materials and Methods. In experiment 5, the histidinol concentration was lowered from that indicated in Fig. 4, because it was found that the presence of G418 and 6-TG in the medium enhanced the effectiveness of histidinol. In experiments 1 to 4, TGR-1 cells were used. In experiments 5 and 6 a *neo*-targeted c-*myc* heterozygote expressing a c-*myc* transgene was used; G418 was included in the selection at 600 μg/ml to prevent the retargeting of the *neo*-targeted gene copy. In all targeting experiments, clonal cell lines were individually analyzed by Southern blotting hybridization. *neo**, mutant *neo*.

flanking probes. Additional nonhomologous insertions were not detected in the homologously targeted clones. Heterozygous cells (cell line HET15 [32]) were plated in increasing concentrations of G418 and were found to be resistant up to 20 mg of G418 per ml (45 mM) (Fig. 3A). This was a surprising result because the c-*myc* gene does not have a strong promoter. We then tested the G418 resistance of 10 clonal nonhomologous cell lines recovered in the targeting experiment. All were strikingly less G418 resistant than homologously targeted cells (Fig. 3B).

The low Neo expression level found in nonhomologously recombined cell lines is not a property of the particular vector construct, or of promoterless vectors in general, but rather is a consequence of the nonhomologous integration process. This was shown in an experiment in which commonly used Neo-expressing vectors (pSV2*neo* and pRSV*neo*) were electroporated into TGR-1 cells and selected in increasing concentrations of G418 (Fig. 3D). The selections were done on pooled populations (as opposed to isolated clonal cell lines) and show that 80% of randomly integrated nonhomologous recombinants are eliminated by increasing G418 concentrations from 600 μg/ml to 5 mg/ml.

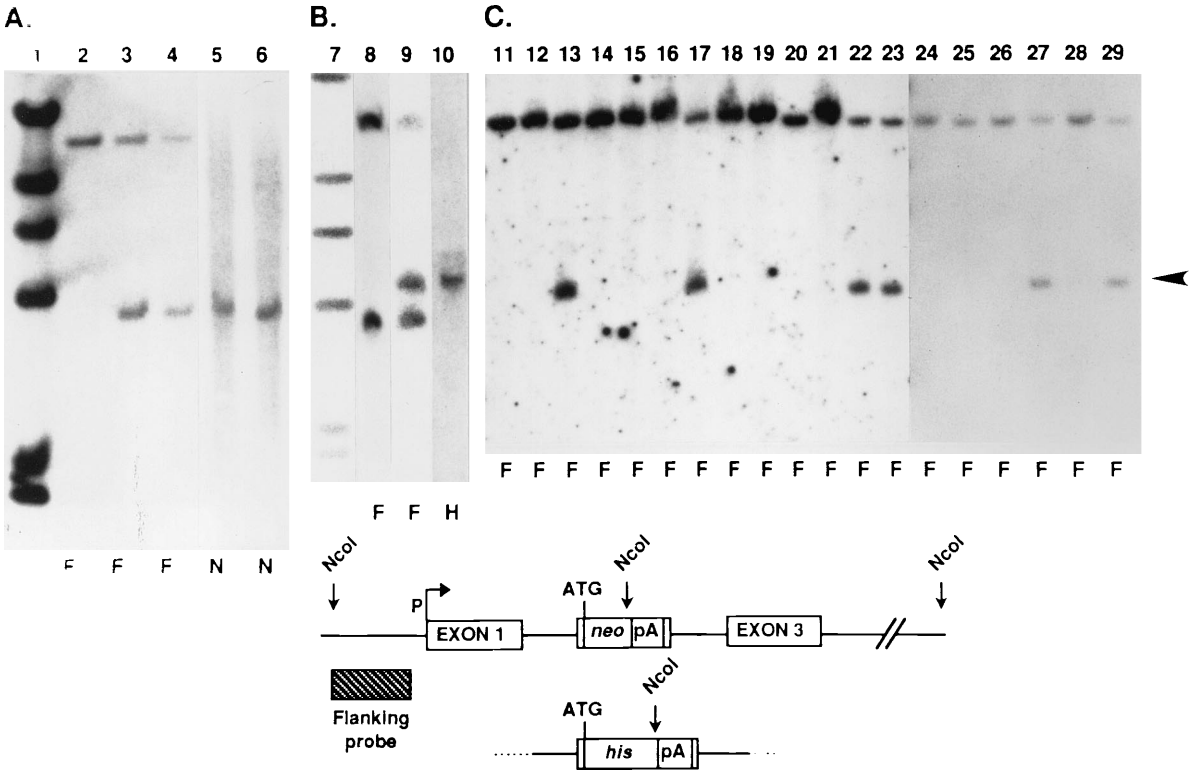

FIG. 2. Analysis of recombination events by Southern blotting hybridization. All samples were digested with *Nco*I (see diagram). The untargeted locus produces an *Nco*I fragment of approximately 15 kbp. Since both the *neo* and *his* sequences contain an *Nco*I site, *neo*- and *his*-targeted loci produce bands of 4.0 and 5.0 kbp, respectively. The probe used (F, 5′ flanking chromosomal probe; N, *neo* probe; H, *his* probe) is indicated below each lane. (A) Targeting of the first copy of c-*myc* with a *neo* vector. Lanes: 1, phage lambda DNA digested with *Hin*dIII; 2, TGR-1 (parental diploid) cell line; 3 and 4, two independent cell lines targeted with a wild-type *neo* vector (Table 1, experiment 2); 5 and 6, same as lanes 3 and 4 but probed with a *neo* probe. (B) Targeting of the second copy of c-*myc* with a *his* vector. Lanes: 7, phage lambda DNA digested with *Hin*dIII; 8, the parental *neo*-targeted heterozygous cell line (Table 1, experiment 3); 9, a *his*-targeted cell line (Table 1, experiment 5); 10, same as lane 9 but probed with a *his* probe. (C) High-efficiency targeting with a mutant *neo* vector. The parental cell line was TGR-1. All of the G418-resistant clones recovered in one experiment (Table 1, experiment 4) are shown. The position of the targeted band is indicated by the arrowhead on the right. Targeted hits are in lanes 13, 17, 22, 23, 27, and 29. pA, polyadenylation site.

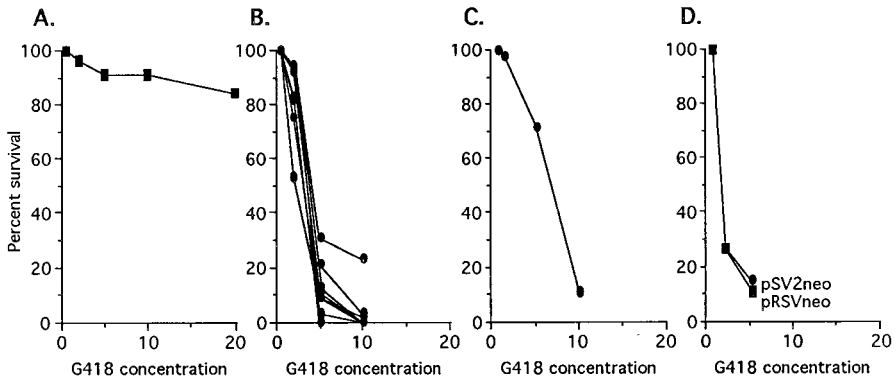

FIG. 3. G418 resistance measured by an efficiency-of-plating assay. (A) Heterozygous c-*myc* cells targeted with a promoterless wild-type *neo* vector; (B) cell lines with nonhomologous insertions of a promoterless wild-type *neo* vector; (C) heterozygous c-*myc* cells targeted with a promoterless mutant-*neo* construct; (D) TGR-1 cells electroporated with pSV2*neo* and pRSV*neo* vectors. Clonal cell lines were used in panels A to C. In panel D, the pooled population following electroporation was used. Cells were trypsinized, diluted, and plated with the indicated concentrations (milligrams per milliliter) of G418. A wide range of dilutions was plated. The medium was buffered with *N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid (HEPES) at pH 7.4. After 14 days at 37°C (feeding every 4 days), dishes were stained with crystal violet and colonies were counted.

Increasing the stringency of selection results in high targeting efficiency. Since the wild-type *neo* gene behaved as an extremely and, in the case of c-*myc*, an unnecessarily powerful selectable marker, we reduced the enzymatic activity of the Neo protein by introducing a point mutation into the *neo* coding region (39) in the promoterless vector. In a repeat of the targeting experiment, 173 colonies were screened and 16 targeted hits were recovered (9% efficiency; Table 1). The original and mutant-*neo* vectors are identical with the exception of the point mutation, but under identical selection conditions (600 μg of G148 per ml and 6 μg of 6-TG per ml), the latter is more efficient. As expected, cell lines targeted with the mutant-*neo* construct displayed increased G418 sensitivity (Fig. 3C). To determine if targeting efficiencies could be further improved, we performed another experiment with the mutant-*neo* vector which was selected at 3 mg of G418 per ml. Only 19 colonies were recovered, but 6 of these were targeted hits (32% efficiency; Fig. 2C; Table 1).

In a parallel selection at 5 mg of G418 per ml, only four colonies were recovered, and one contained a targeted event. The targeted clone grew very slowly at 5 mg of G418 per ml (but grew well at 3 mg/ml), and it is likely that others were lost during the ring cloning of the cell lines. G418 at 3 mg/ml is thus the maximum selection stringency for c-*myc*, since the c-*myc* promoter is apparently too weak to express enough of the mutant *neo* gene for effective growth above this concentration of G418. This finding, in turn, indicates that efficiencies above 32% should be possible with genes that have promoters stronger than c-*myc*.

We estimated, using quantitative Western blotting analysis, that homologously targeted cells produce 5×10^4 to 1×10^5 Neo molecules per cell (approximately 0.001% of total cellular protein [16]). This level of expression if typical of low-abundance housekeeping genes, such as the adenylyl cyclase (0.003% [26]), SP1 transcription factor (0.003% [20]), cdc2 kinase (0.02% [8]), and protein kinase C (0.1% [19]) genes. A nonhomologously recombined clone with an efficiency of plating of 5% at 2 mg of G418 per ml did not display a detectable Neo signal on Western blots. We estimated that the detection limit was approximately 5×10^3 Neo molecules per cell.

Selectable genes for multiple rounds of gene targeting. One approach to generating homozygously disrupted cell lines is to perform a second round of targeting, which requires a vector with a distinct positively selectable gene. Experiments with *neo* vectors demonstrated the importance of two aspects of the promoterless strategy: (i) the level of marker gene expression necessary to generate a drug-resistant clone and (ii) the expression levels elicited by individual (nonhomologous as well as homologous) recombinant events. We applied a promoter trap method (14, 15, 36) to assess the efficacy of positive markers and their suitability for the promoterless strategy. In an experiment of this design, the fragments are randomly probing the genome for sites that allow sufficient expression to yield drug-resistant clones. Promoterless fragments with a wild-type *neo*, mutant *neo*, *hyg*, or *his* open reading frame were excised from the corresponding targeting vectors (Fig. 1) and electroporated into cells, and drug-resistant clones were selected. The fragments were isogenic with the exception of the marker sequences and contained, in the following order, a short 3′ segment of c-*myc* intron 1, the beginning of exon 2, a selectable gene open reading frame, the simian virus 40 late polyadenylation site, the end of c-*myc* exon 2, and a short 5′ segment of c-*myc* intron 2.

The results of gene trap experiments showed that wild-type *neo* is the most powerful selectable marker, mutant-*neo* and *his* are intermediate, and *hyg* is very weak (Fig. 4). This finding, in turn, predicts that a *his* vector should target c-*myc* with an efficiency approximately that of the mutant-*neo* vector and that a *hyg* vector may not be able target c-*myc*. When the *his* and *hyg* isogenic vectors were used in targeting experiments, the *his* vector yielded seven homologous hits from a total of 23 clones (30% efficiency; Fig. 2B; Table 1), which is very similar to the efficiency obtained with the mutant-*neo* vector (32%), whereas the *hyg* vector yielded no homologous hits from a total of 48 clones. In both cases, the recipient cell line was a *neo*-targeted c-*myc* heterozygote (cell line HET40) expressing a c-*myc* transgene (Myc-estrogen receptor chimera [13]), and the cells were grown in the presence of estrogen in the targeting experiments. Under these conditions, the HET40 cell line grows at the same rate as the parental TGR-1 cell line. The promoter trap experiment thus accurately predicted the targeting frequencies.

The following trivial explanations for the failure of the *hyg* vector to target are unlikely: (i) position effects (*hyg* was in the same position as *neo* and *his*, the flanking c-*myc* sequences were the same, and the recombinant junction was confirmed by DNA sequencing); (ii) translational effects (the Kozak consensus sequence surrounding the ATG initiation codon was the same in *neo*, *his*, and *hyg* constructs); (iii) stringency of selec-

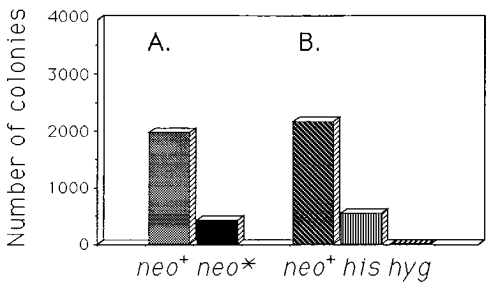

FIG. 4. Estimation of selectable marker efficacy in gene trap experiments. Parts A and B show experiments performed on separate occasions under identical conditions. Restriction fragments containing the indicated selectable markers were excised from the corresponding targeting vectors (Fig. 1) and gel purified. Recovery was determined by gel electrophoresis of dilution series and comparison with standards of known concentration. Equivalent amounts of DNA (5 μg) were electroporated in parallel into TGR-1 cells. Selection conditions: wild-type *neo* (*neo*$^+$) and mutant-*neo* (*neo**), 600 μg of G418 per ml; *his*, 86 μg of histidinol per ml (0.4 mM); *hyg*, 110 μg of hygromycin B per ml. In separate experiments, 110 μg of hygromycin B per ml was determined to be the minimum concentration required to effectively eliminate sensitive TGR-1 cells. A wide range of dilutions was plated. After 14 days at 37°C (feeding every 4 days), dishes were stained and colonies were counted.

tion (the drug concentration used was the minimum necessary [110 μg of hygromycin B per ml] to eliminate sensitive cells); and (iv) mutations in *hyg* sequences (the *hyg* gene was obtained from a common vector [pY3 {3}] and functioned well in that plasmid or when recloned with another promoter). In addition, a plasmid in which the c-*myc* promoter was reconstructed into the targeting vector yielded very few Hyg-resistant clones, whereas insertion of a long terminal repeat promoter restored the yield of Hyg-resistant clones to the level of the original pY3 plasmid. We conclude that the *hyg* vector failed to target because the c-*myc* promoter is too weak to express sufficient Hyg protein to generate a drug-resistant clone.

Homogenotization of heterozygously targeted loci. Two strategies can be used to obtain homozygous cell lines: direct targeting of the second gene copy and homogenotization of a heterozygously targeted clone by selection in high drug concentrations (24). The probable mechanism of homogenotization is duplication of the targeted chromosome in combination with loss of the untargeted chromosome. Non-sister chromatid exchange is such an exceedingly rare event in mitotic cells that gene conversion is an unlikely mechanism (4). Mitotic nondisjunction, which can lead to either chromosome duplication or loss, has been documented to occur at frequencies of 10^{-3} to 10^{-4} in many cell lines (6, 7, 33). We selected mutant-*neo*-targeted cells (Fig. 3C) at 20 mg of G418 per ml (efficiency of plating of 1 to 2%), but among 51 cell lines recovered, no homozygotes were found. Lethality of the Myc$^-$ phenotype cannot account for the failure to homogenotize, because the cells were transfected with a conditional Myc-expressing transgene (13).

We examined the propensity for nondisjunction by introducing single-copy insertions of the *gpt* gene and testing individual clones for spontaneous resistance to 6-TG. Single-copy *gpt* integrants were obtained by both infection with the retroviral vector pZIP*gpt* and electroporation with the plasmid vector pSV2*gpt*. Clonal cell lines were established and screened by Southern hybridization, and cell lines with multiple integration events were discarded. Fifteen single-copy Gpt$^+$ cell lines (nine ZIP*gpt* derived and six pSV2*gpt* derived) were tested for reversion to a Gpt$^-$ phenotype be selection in 6-TG (Materials and Methods). Two distinct frequency classes were observed: high (10^{-3} to 10^{-4}) and low (10^{-6} to 10^{-7}). Four cell lines fell into the high-frequency class (2 each ZIP*gpt* and pSV2*gpt* derived), and 11 cell lines fell into the low-frequency class (7 ZIP*gpt* derived and 4 pSV2*gpt* derived). Selections repeated in an independent experiment with three cell lines produced the same results.

Only the low-frequency class is consistent with spontaneous mutagenesis of the *gpt* gene in several cell lines (2) as well as with spontaneous mutagenesis of the *hprt* gene in Rat-1 cells (27). Southern blotting hybridization showed that a *gpt* signal was lost in cell lines of the high-frequency class but not of the low-frequency class. Reversion to 6-TG resistance in the high-frequency class could be enhanced approximately 10-fold by sublethal concentrations of colcemid, a treatment known to increase chromosome nondisjunction. These results indicate that the Gpt$^-$ phenotype was due to point mutagenesis in the low-frequency class and to mitotic nondisjunction in the high-frequency class. The failure to homogenotize is thus not a general property of the TGR-1 cell line, and, most importantly, some chromosomes apparently cannot be homogenotized by nondisjunction.

DISCUSSION

A systematic evaluation of the biological parameters operative during the selection process for homologously targeted events has not been reported; investigators have typically employed minimum drug concentrations required to elicit effective killing, and various selectable markers have been used essentially interchangeably. We found that enrichment ratios of up to 5,000- to 10,000-fold can be obtained with promoterless vectors by exploiting the biological characteristics of selectable markers and the chromosomal target gene. Our examination of commonly used selectable genes revealed that widely differing levels of expression are necessary to generate drug-resistant clones. A powerful selectable marker is required to target a gene expressed at a low level, whereas weaker markers can be used with more abundantly expressed genes. A powerful marker, however, also generates a high background of nonhomologous events, since many exceed the low expression threshold required to generate drug-resistant clones. We also found that marker genes integrated homologously are expressed at a significantly higher level than those integrated nonhomologously, which provides a large window of opportunity for differential selection.

The high efficiencies that we observed are not due to an unusual recombinogenicity of the c-*myc* locus, since a conventional isogenic PNS vector had an efficiency of less than 0.5%. The high efficiencies are due to the progressive elimination of nonhomologous recombinants. This is evidenced by the fact that the absolute frequencies of homologous recombination were essentially identical in all our experiments: 5.9×10^{-6} (one hit in 1.7×10^5 cells surviving electroporation) in the initial experiment (Table 1, experiment 2; 4% efficiency) and 3×10^{-6} (one hit in 3.3×10^5 surviving cells) in the final high-efficiency experiment (Table 1, experiment 4; Fig. 2C; 32% efficiency). A comparison of these absolute frequencies with frequencies typically obtained in ES cells (37) also underscores the fact that the c-*myc* locus in TGR-1 cells is actually of relatively low recombinogenicity. The low frequency of targeting is not due to inefficient electroporation, since the absolute frequencies of both transient and stable transfection (2.5×10^{-1} and 1×10^{-2} to 3×10^{-2}, respectively) are quite high in TGR-1 cells (27). From these values, the overall enrichment ratio can be calculated to be 5×10^3 to 1×10^4.

The level at which a selectable gene must be expressed to generate a drug-resistant clone, a property that we refer to as

marker efficacy, is an important parameter in gene targeting. We have used a biological promoter trap assay to assess the efficacy of several selectable markers and found large differences among them. The *hyg* gene produced 200-fold-fewer colonies than wild-type *neo*. This finding indicates that wild-type *neo* is a powerful selectable marker and accounts for the fact that it can be used to target a gene with a relatively weak promoter, such as c-*myc* (albeit with intermediate efficiency, since nonhomologous hits with low expression can also score). This finding also predicts that wild-type *neo* can be used to target genes with even weaker promoters, which has been documented (10, 22). *hyg* scores as a very weak selectable marker, which explains why it cannot be used to target a gene like c-*myc*. This finding predicts that *hyg* can be used to efficiently target genes with strong promoters, which has also been documented (35). The mutant-*neo* and *his* genes are intermediate: they can be adequately expressed from the c-*myc* promoter, but since they are weaker than wild-type *neo*, better efficiencies of targeting are obtained. In all cases examined, the results of promoter trap experiments agreed with subsequent targeting efficiencies. The efficacies of selectable markers can be fine-tuned by adjusting the drug concentrations in the medium.

Since the selection strategy that we have developed is based on exploiting the biological properties of selectable genes, it should be generally applicable to other loci as well as cell lines. In other words, whatever the recombinogenicity of a given chromosomal locus, our strategy should achieve a high enrichment ratio on top of that intrinsic frequency. The selection strategy is based primarily on the surprising observation that nonhomologous recombination following electroporation results in poor expression of the transfected genes. Since the enrichment ratio is determined by both the expression level of the target gene (the lower the expression, the lower the enrichment) and its intrinsic recombinogenicity, we believe that a significant point in assessing the general applicability of the strategy is that the c-*myc* gene is a good representative of an "average" gene likely to be encountered in future targeting experiments. In other words, the c-*myc* gene is not expressed at a high level, and it is not unusually recombinogenic.

Several strategies can be used to boost the expression level of the selectable marker protein (and thus the overall enrichment) from very weak target promoters. In the case of c-*myc*, the mRNA has a short half-life; placing a polyadenylation site immediately 3′ to the selectable marker removed the major destabilization sequence from the mRNA and greatly enhanced expression. Many low-abundance proteins are translated from messages with poor Kozak box consensus sequences; replacement with an optimal sequence would enhance expression of the selectable marker protein. Our strategy of having the c-*myc* start codon synonymous with that of the selectable marker would allow this strategy and also precludes problems with multiple translational starts or reduced enzymatic activity of fusion proteins. This strategy, however, has the disadvantage in cases where the translation start and the promoter are separated by insufficient length of sequence to construct an effective targeting vector.

As a method to obtain homozygous knockout cell lines, the successful use of a homogenotization method has been documented in a number of cases (5, 24), but we as well as others (1) have found that it is not universally applicable. We believe that this is because homogenotization occurs by mitotic chromosome nondisjunction, and the loss of certain chromosomes would be prevented by the presence of recessive lethal mutations. The acquisition of recessive lethal mutations may be a general property of cell lines that have been extensively cultured in vitro. Another disadvantage of homogenotization is that unless cells are recently derived from highly inbred animals, the nondisjunction derivative and the parental cell line cannot be considered isogenic. To directly target two gene copies, targeting vectors with distinct selectable markers are required. In this context, it is significant that the *his* marker displayed the same high efficiency as the mutant-*neo* marker. We have thus developed the tools to rapidly and efficiently target two gene copies in any diploid cell line.

Our data on Neo expression in c-*myc* recombinants, and the relative efficacies of selectable markers, can be used as a guide in the design of future targeting strategies. We recommend the following procedure to optimize selection conditions (28). The expression level of the target gene is quantitated in the cell line of interest, and this information is used to choose a selectable marker. A targeting vector incorporating the marker gene is constructed, cells are electroporated, and aliquots are selected with increasing drug concentrations. For each aliquot, plates are trypsinized to establish several pools (without ring cloning of colonies), and each pool is analyzed by PCR. Targeted events are typically present in pools selected at low drug concentrations. As the concentration of drug at which the target gene promoter is unable to sufficiently express the selectable marker is reached, homologous recombinants rapidly disappear. Total colony numbers (the sum of homologous and illegitimate events) decrease continuously with increasing drug concentrations. The objective of this pilot experiment is to determine the maximum drug concentration at which targeted events can be recovered. Targeting efficiencies can also be calculated. The electroporation is then repeated, with selection at the drug concentration determined to yield the maximum efficiency of targeting, and the requisite number of individual colonies are cloned.

ACKNOWLEDGMENTS

We thank Steve Prouty and Michelle Follansbee for contributing unpublished data. We gratefully acknowledge Steve Prouty, Peter Lengyel, and Mark Solomon for helpful discussions throughout the course of this work. We thank Charles Radding and Alison DeLong for substantial advice during the writing of the manuscript.

This work was supported by NIH grant GM-R01-41690 and Presidential Young Investigator Award DMB-9057715 from the National Science Foundation to J.M.S. K.D.H. was supported in part by NIH training grant GM 07223.

REFERENCES

1. **Arbones, M. L., H. A. Austin, D. J. Capon, and G. Greenburg.** 1994. Gene targeting in normal somatic cells: inactivation of the interferon-γ receptor in myoblasts. Nat. Genet. **6:**90–97.
2. **Ashman, C. R., and R. L. Davidson.** 1987. Sequence analysis of spontaneous mutations in a shuttle vector gene integrated into mammalian chromosomal DNA. Proc. Natl. Acad. Sci. USA **84:**3354–3358.
3. **Blocklinger, K., and H. Diggelmann.** 1984. Hygromycin B phosphotransferase as a selectable marker for DNA transfer experiments with higher eucaryotic cells. Mol. Cell. Biol. **4:**2929–2931.
4. **Bollag, R. J., A. S. Waldman, and R. M. Liskay.** 1989. Homologous recombination in mammalian cells. Annu. Rev. Genet. **23:**199–225.
5. **Boylan, J. F., D. Lohnes, R. Taneja, P. Chambon, and L. J. Gudas.** 1993. Loss of retinoic acid receptor gamma function in F9 cells by gene disruption results in aberrant Hoxa-1 expression and differentiation upon retinoic acid treatment. Proc. Natl. Acad. Sci. USA **90:**9601–9605.
6. **Chasin, L. A.** 1973. The effect of ploidy on chemical mutagenesis in cultured chinese hamster cells. J. Cell. Physiol. **82:**299–307.
7. **Chasin, L. A., and G. Urlaub.** 1976. Mutant alleles for hypoxanthine phosphoribosyltransferase: codominant expression, complementation, and segregation in hybrid Chinese hamster cells. Somatic Cell. Genet. **2:**453–467.
8. **Cisek, L. J., and J. L. Corden.** 1991. Purification of protein kinases that phosphorylate the repetitive carboxyl-terminal domain of eukaryotic RNA polymerase II. Methods Enzymol. **200:**301–325.
9. **Cone, R. D., E. B. Reilly, H. N. Eisen, and R. C. Mulligan.** 1987. Tissue-specific expression of functionally rearranged lambda 1 Ig gene through a

retrovirus vector. Science **236:**954–957.
10. **DeGregori, J. V., A. Russ, H. von Melchner, H. Rayburn, P. Priyaranjan, N. A. Jenkins, N. G. Copeland, and H. E. Ruley.** 1993. A murine homolog of the yeast RNA1 gene is required for postimplantation development. Genes Dev. **8:**256–276.
11. **Doetschman, T., N. Maeda, and O. Smithies.** 1988. Targeted mutation of the Hprt gene in mouse embryonic stem cells. Proc. Natl. Acad. Sci. USA **85:**8583–8587.
12. **Dorin, J. R., J. D. Inglis, and D. J. Porteous.** 1989. Selection for precise chromosomal targeting of a dominant marker by homologous recombination. Science **243:**1357–1360.
13. **Eilers, M., D. Picard, K. R. Yamamoto, and J. M. Bishop.** 1989. Chimaeras of myc oncoprotein and steroid receptors cause hormone-dependent transformation of cells. Nature (London) **340:**66–68.
14. **Friedrich, G., and P. Soriano.** 1991. Promoter traps in embryonic stem cells: a genetic screen to identify and mutate developmental genes in mice. Genes Dev. **5:**1513–1523.
15. **Gossler, A., A. L. Joyner, J. Rossant, and W. C. Skarnes.** 1989. Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes. Science **244:**463–465.
16. **Hanson, K. D., and J. M. Sedivy.** Unpublished observations.
17. **Hanson, K. D., M. Shichiri, M. R. Follansbee, and J. M. Sedivy.** 1994. Effects of c-*myc* expression on cell cycle progression. Mol. Cell. Biol. **14:**5748–5755.
18. **Hartman, S. C., and R. C. Mulligan.** 1988. Two dominant-acting selectable markers for gene transfer studies in mammalian cells. Proc. Natl. Acad. Sci. USA **85:**8047–8051.
19. **Huang, K. P., and F. L. Huang.** 1991. Purification and analysis of protein kinase C isozymes. Methods Enzymol. **200:**241–252.
20. **Jackson, S. P., and R. Tjian.** 1989. Purification and analysis of RNA polymerase II transcription factors by using wheat germ agglutinin affinity chromatography. Proc. Natl. Acad. Sci. USA **86:**1781–1785.
21. **Jasin, M., and P. Berg.** 1988. Homologous integration in mammalian cells without target gene selection. Genes Dev. **2:**1353–1363.
22. **Jeannotte, L., J. C. Ruiz, and E. J. Robertson.** 1991. Low level of Hox1.3 gene expression does not preclude the use of promoterless vectors to generate a targeted gene disruption. Mol. Cell. Biol. **11:**5578–5585.
23. **Mansour, S. L., K. R. Thomas, and M. R. Capecchi.** 1988. Disruption of the proto-oncogene int-2 in mouse embryo-derived stem cells: a general strategy for targeting mutations to non-selectable genes. Nature (London) **336:**348–352.
24. **Mortensen, R. C., D. A. Conner, S. Chao, A. A. T. Geisterfer-Lowrance, and J. G. Seidman.** 1992. Production of homozygous mutant ES cells with a single targeting construct. Mol. Cell. Biol. **12:**2391–2395.
25. **Mulligan, R. C., and P. Berg.** 1981. Selection for animal cells that express the Escherichia coli gene coding for xanthine-guanine phosphoribosyltransferase. Proc. Natl. Acad. Sci. USA **78:**2072–2076.
26. **Pfeuffer, E., S. Mollner, and P. Pfeuffer.** 1991. Purification of adenylyl cyclase from heart and brain. Methods Enzymol. **195:**83–91.
27. **Prouty, S. M., K. D. Hanson, A. L. Boyle, J. R. Brown, M. Shichiri, M. R. Follansbee, W. Kang, and J. M. Sedivy.** 1993. A cell culture model system for genetic analyses of the cell cycle by targeted homologous recombination. Oncogene **8:**899–907.
28. **Prouty, S. M., and J. M. Sedivy.** Unpublished observations.
29. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
30. **Sedivy, J. M., and A. L. Joyner.** 1992. Gene targeting. W. H. Freeman and Co., New York.
31. **Sedivy, J. M., and P. A. Sharp.** 1989. Positive genetic selection for gene disruption in mammalian cells by homologous recombination. Proc. Natl. Acad. Sci. USA **86:**227–231.
32. **Shichiri, M., K. D. Hanson, and J. M. Sedivy.** 1993. The effects of c-myc expression on proliferation, quiescence, and the G_0 to G_1 transition in nontransformed cells. Cell Growth Differ. **4:**93–104.
33. **Siminovitch, L.** 1976. On the nature of hereditable variation in cultured somatic cells. Cell **7:**1–11.
34. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region. J. Mol. Appl. Genet. **1:**327–341.
35. **te Riele, H., E. R. Maandag, A. Clarke, M. Hooper, and A. Berns.** 1990. Consecutive inactivation of both alleles of the pim-1 proto-oncogene by homologous recombination in embryonic stem cells. Nature (London) **348:**649–651.
36. **von Melchner, H., J. V. DeGregori, H. Rayburn, S. Reddy, C. Friedel, and H. E. Ruley.** 1992. Selective disruption of genes expressed in totipotent embryonal stem cells. Genes Dev. **6:**919–927.
37. **Waldman, A. S.** 1992. Targeted homologous recombination in mammalian cells. Crit. Rev. Oncol. Hematol. **12:**49–64.
38. **Yagi, T., Y. Ikawa, K. Yoshida, Y. Shigetani, N. Takeda, I. Mabuchi, T. Yamamoto, and S. Aizawa.** 1990. Homologous recombination at c-fyn locus of mouse embryonic stem cells with use of diptheria toxin A-fragment gene in negative selection. Proc. Natl. Acad. Sci. USA **87:**9918–9922.
39. **Yenofsky, R. L., M. Fine, and J. W. Pellow.** 1990. A mutant neomycin phosphotransferase II gene reduces the resistance of transformants to antibiotic selection pressure. Proc. Natl. Acad. Sci. USA **87:**3435–3439.