Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 8067–8071, September 1991
Genetics

Fidelity of targeted recombination in human fibroblasts and murine embryonic stem cells

(homologous recombination/mutation/chemical cleavage/hypoxanthine phosphoribosyltransferase)

HUI ZHENG*, PAUL HASTY*, MARK A. BRENNEMAN*, MARKUS GROMPE*, RICHARD A. GIBBS*, JOHN H. WILSON*†, AND ALLAN BRADLEY*‡

*Institute for Molecular Genetics and †Verna and Marrs McLean Department of Biochemistry, Baylor College of Medicine, Houston, TX 77030

Communicated by Ralph L. Brinster, May 20, 1991

ABSTRACT **Targeted recombination in murine embryonic stem cells promises to be a powerful tool for introducing specific mutations into target genes to study development in mice and to create animal models of human disease. Gene targeting also holds potential for correcting genetic defects as an approach to human gene therapy. To precisely modify target genes, homologous recombination must proceed with high fidelity. However, several results have suggested that targeted recombination may be highly mutagenic. To test the accuracy of gene targeting we analyzed 44 independent targeted recombinants at the hypoxanthine phosphoribosyltransferase (HPRT) locus in a human fibroblast cell line and in mouse embryonic stem cells. We surveyed 80 kilobases around the sites of recombination by using chemical cleavage of mismatches. Only two mutations were found: a T → G transversion and a thymidine deletion. Thus, gene targeting in mammalian cells can be extremely accurate. These results demonstrate the feasibility of generating precise modifications of mammalian genomes by gene targeting.**

Gene targeting has become a powerful tool for functional analysis of genes *in vivo*. In conjunction with mouse embryonic stem (ES) cell technology, it has made it possible to generate mice with mutations at virtually any desired locus (1–4). Much of the work to date using gene targeting has involved insertion of a selectable marker into an exon of the target gene, so that the reading frame is interrupted. Although this approach is valuable to study the effects of gene ablation, detailed analysis of a gene product and the control of its expression will require modification at the level of specific nucleotides. To achieve such precise modifications, targeted recombination must proceed with very high fidelity.

In contrast to intrachromosomal recombination, which seems very accurate in mammalian cells (5), several lines of evidence have suggested that targeted recombination may induce additional mutations with high frequency (6–8). The first evidence of induced mutations came from experiments by Thomas and Capecchi (6), in which a deficient neomycin resistance (neo) gene was corrected by recombination. Of the cell colonies with restored neo function, fully 50% resulted from the acquisition of new mutations in the neo gene, rather than from correction of the original mutation. The actual mutation frequency may have been even higher because mutations that did not generate a functional neo gene would not have been scored. Doetschman *et al.* (7) have reported an experiment in which two of six targeted recombination events at the hypoxanthine phosphoribosyltransferase (HPRT) locus in mouse ES cells were accompanied by deletion mutations. A particularly striking instance has been reported by Brinster *et al.* (8). In gene targeting designed to correct a deficient mouse major histocompatibility complex class II $E\alpha$ gene directly in mice by pronuclear injection, 15 new mutations were introduced into a segment of the target gene only 1.5 kilobases (kb) long.

In this work, we have addressed the question of fidelity by analyzing targeted recombination events at the HPRT locus in two cell types: the immortalized human fibroblast line HT1080 (9), and mouse ES cells (1). We have used a strategy of chemical cleavage of mismatches (10, 11), which has permitted us to scan large stretches of DNA in the vicinity of recombination events for mutation. The results show that targeted recombination in both cell types is extremely accurate. A comparison of our results with previous data suggests conditions under which precise modifications of mammalian genomes can be accomplished by gene targeting.

MATERIALS AND METHODS

Electroporation and Selection of Targeted HPRT Recombinants. The targeted human HPRT clones were derived by electroporating the *Hin*dIII-digested vector RV6.9h (Fig. 1*A*) into the HT1080 cell line, by using a Bio-Rad Gene Pulser apparatus. Cells (1×10^7) were resuspended in 800 μl of phosphate-buffered saline containing 50 μg of DNA and electroporated at 250 μF and 1125 V/cm. The treated cells were selected in G418 at 400 μg/ml 24 hr later, and 6-thioguanine at 5 μg/ml was applied 10 days after G418 selection. Independent G418 and 6-thioguanine-resistant colonies were picked, grown up, and analyzed by PCR and/or Southern analysis. The maintenance of ES cultures and the targeting process on the mouse HPRT gene in AB1 ES cells have been described (1, 12).

Chemical Cleavage of Mismatches. The chemical mismatch cleavage reaction was done essentially as described, except that 250 ng of genomic DNA was used as template for PCR, and the hydroxylamine reaction was shortened to 10–15 min (11). The products were resolved by electrophoresis in a 4% denaturing polyacrylamide gel followed by autoradiography.

PCR Direct Sequencing. The method for PCR direct sequencing has been described (13). For sequencing of the human HPRT mutant T11, the template was obtained by a PCR using primers b and d to amplify the downstream fragment containing the cleavage site (Fig. 3*B*). A single-strand PCR reaction was done by using the first PCR product and primer d only. Primer b was labeled with [γ-^{32}P]ATP through a kinase reaction and used as the annealing primer for dideoxynucleotide chain-termination sequencing.

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; neo, neomycin resistance; HPRT, hypoxanthine phosphoribosyltransferase.

‡To whom reprint requests should be addressed.

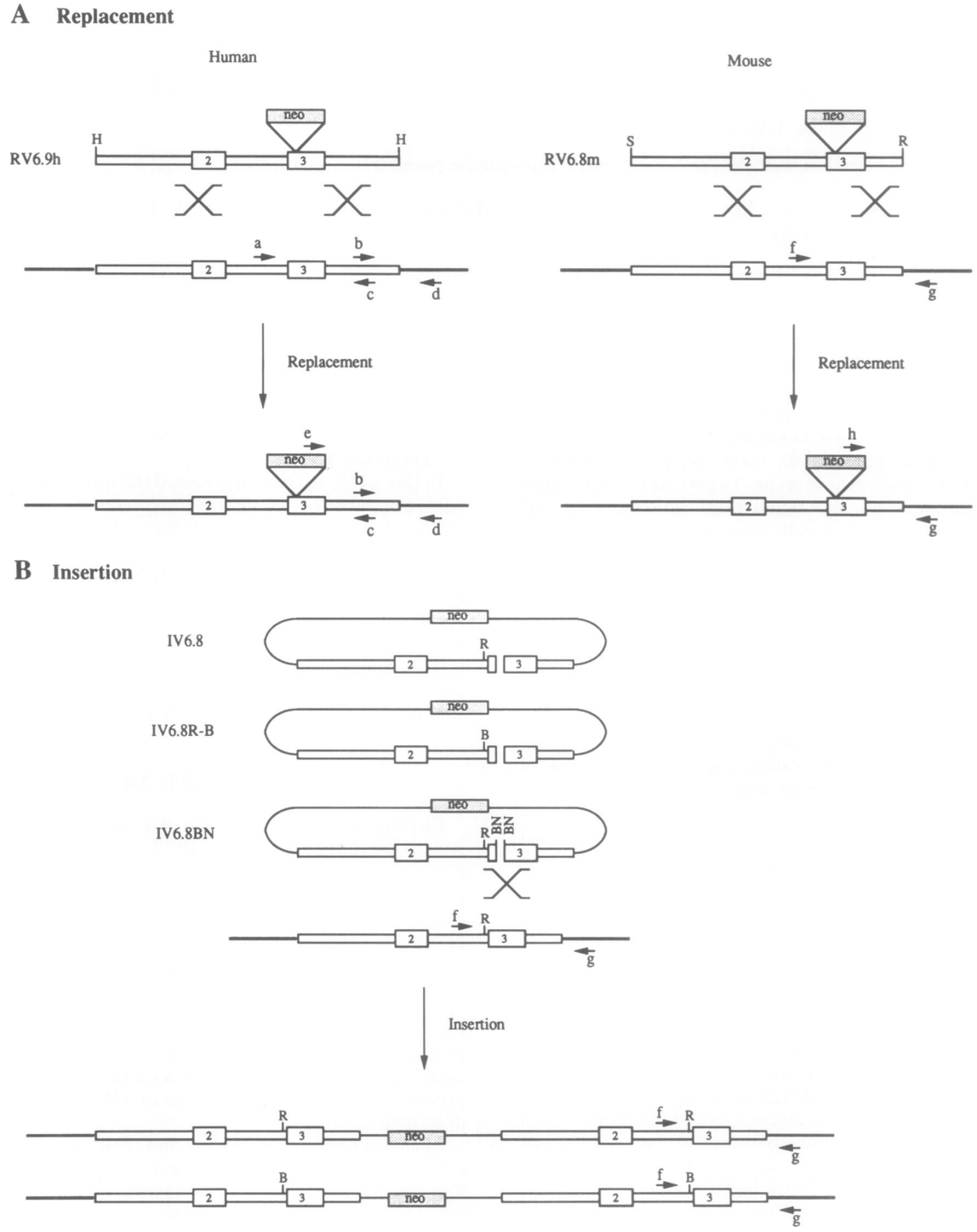

FIG. 1. Gene targeting pathway and primer design for mutation detection. (*A*) Gene targeting at the HPRT locus using replacement vectors. The 2.1-kb arm of the human gene targeting vector was separated into two PCR products for chemical mismatch cleavage reaction. Primers a and c and e and c amplified the upstream fragment of wild-type (1.7 kb) and targeted clones (1.5 kb), respectively, and primers b and d amplified the downstream fragment 1.5 kb in length. Primers b and c were overlapped by 150-bp sequences. The 1.2-kb 3′ arm of the mouse vector RV6.8m was amplified as a single PCR fragment using primers h and g, which generated a fragment of 1.5 kb. Primers f and g were used to amplify wild-type AB1 cells. The orientation of each primer was shown by the arrow. H, *Hin*dIII; S, *Sac* I; and R, *Eco*RI. Sequences of primers are as follows: a, 5′-CCT TAT GAA ACA TGA GGG CAA AGG-3′; b, 5′-CAG GGT AGA AAT GCT ACT TCA GGC-3′; c, 5′-GCT CGA CAT TGG GTG GAA ACA TTC-3′; d, 5′-TAA AGG CTG GCC AGT TAT ACT GCC-3′; e, 5′-CAC TAC TTC CTA CCC ATA AGC TCC-3′; f, 5′-TTA TGA TTA CAG TCC AGC CTG GGC-3′; g, 5′-CTC TGT GAG CCA CTA TGT AAG TGC-3′; and h, 5′-ATC GCC TTC TAT CGC CTT CTT GAC-3′. (*B*) Gene targeting at the mouse HPRT locus using insertion vectors. Primers f and g were used to amplify both the wild type and the 3′-duplication of the targeted clones. R, *Eco*RI; B, *Bam*HI; and BN is part of the *Bam*HI–*Not* I–*Bam*HI linker derived by linearizing with *Not* I. The oligonucleotide sequences are as follows: *Bam*HI linker, 5′-AAT TGC CGG ATC CGG CCG GAT CCG GGC-3′; *Bam*HI–*Not* I–*Bam*HI linker, 5′-TCG ATG GAT CCG CGG CCG CGG ATC CA-3′.

Subcloning Sequencing. The wild-type AB1 DNA and the mutant IV6.8BN J8 were amplified by primers f and g. The PCR products were digested with *Xho* I and *Eco*RI and subcloned into pTZ18u vector. Automated DNA sequencing was done on an AB1 370A fluorescent sequencer.

RESULTS

Gene Targeting at the Human and Mouse HPRT Locus. For these experiments we examined targeting to the same region of the HPRT gene—≈6.8 kb that encompasses exons 2 and 3—in human HT1080 cells and mouse ES cells. We chose to disrupt the wild-type HPRT gene, rather than correct a defective gene, to avoid selecting against potential recombination-induced mutations. Two kinds of vector architecture were tested: replacement vectors were used in human and mouse cells (Fig. 1*A*), and insertion vectors were used in mouse cells (Fig. 1*B*) (14). The targeting vector for human cells RV6.9h (Fig. 1*A*) consisted of a 6.9-kb *Hin*dIII segment of the HPRT gene (15). A pMC1neopA gene (14) was inserted at the *Xho* I site in exon 3, which separated the HPRT sequence into 4.8-kb 5′ and 2.1-kb 3′ elements. The replacement vector for mouse ES cells RV6.8m (Fig. 1*A*) contained a 6.8-kb *Sac* I–*Eco*RI fragment of the mouse HPRT sequence (16), which was interrupted by insertion of a PGKneobpA gene into the *Xho* I site of exon 3 (2). The lengths of 5′ and 3′ homologous sequences were 5.6 kb and 1.2 kb, respectively.

Three insertion vectors were used: IV6.8, IV6.8R-B, and IV6.8BN (Fig. 1*B*). All of these contained the same HPRT sequence as RV6.8m, the pTZ vector, and the pMC1neopA gene inserted in the plasmid backbone. Vector IV6.8 contained the wild-type HPRT sequence. Vectors IV6.8R-B and IV6.8BN were derived from vector IV6.8 by inserting a *Bam*HI linker into the *Eco*RI site in HPRT intron 2 or a *Bam*HI–*Not* I–*Bam*HI linker into the *Xho* I site of exon 3, respectively. Vectors IV6.8 and IV6.8R-B were linearized at the *Xho* I site, whereas vector IV6.8BN was linearized by *Not* I digestion before being introduced into ES cells. Targeted recombinants the structures of which were verified to match those shown in Fig. 1 were used to assess the accuracy of gene targeting.

Primer Design for PCR Amplification of Targeted Recombinants. In all cases the focus for analysis was the short arm of the targeting vector; in this way we could cover the entire region in which a recombination event occurred. Using PCR technology we amplified fragments across target regions in the parent chromosome and from the targeted recombinants, using primers as indicated in Fig. 1. Because mutations at the ends of vectors had been previously observed (1, 7), we positioned the 3′ primer 500 and 150 nucleotides outside the end of homology with the vectors for targeting human and mouse cells, respectively, so that mutations immediately adjacent to the target site would be detected. In the initial phase of the work, PCR products from targeted recombinants in human cells were sequenced directly. A total of 2 kb of DNA sequence from three independent recombinants were analyzed and were found to match perfectly with the published sequence (data not shown). These results suggested that the mutation rate may be low in human cells and that we would need to compare larger stretches to find evidence for mutations.

Chemical Cleavage of Mismatches. To compare large stretches of sequence rapidly, we used the technique of chemical cleavage of mismatches (10, 11). In this method, PCR fragments from the wild-type and targeted clones are generated, and heteroduplexes are created between the two PCR products. Any sequence difference leads to a mismatch that must expose either a thymidine or a deoxycytidine in one of the two possible heteroduplexes (Fig. 3). Thymidines and deoxycytidines in mismatches are susceptible to modification by either hydroxylamine (which reacts with deoxycytidines) or by osmium tetroxide (which reacts primarily with thymidines and to a lesser extent with deoxycytidines). Subsequent cleavage by piperidine generates fragments the lengths of which correspond to the sites of the mismatches. By using a single-labeled primer in the PCR reactions, it is possible to identify individual strands in either heteroduplex and, therefore, to detect all mutations (11). Because the frequency of mutations was apparently low, we set up several positive control mismatches for the cleavage reactions from known mutants as indicated in Figs. 2 and 3 (17). It was shown that all three classes of mutations—i.e., point mutations, deletions, and insertional mutations, could be correctly detected.

Analysis of Clones Targeted with Replacement Vectors. Using chemical cleavage strategy, we analyzed 11 human clones that arose from targeting with the replacement vector RV6.9h (Fig. 1*A*). In >30 kb screened, a sequence difference was detected in only one recombinant (Fig. 3*B*). Direct sequencing of the PCR product revealed a T → G transversion mutation at the site of cleavage (Fig. 3*B*). In this targeted recombinant, the wild-type sequence was also evident in the sequencing gel, suggesting a mixed colony (which could have resulted either from an error of replication that occurred after a single targeted event or from two separate targeted events, only one of which was accompanied by a mutation). The presence of only one mutation in the 30 kb around the targeting sites suggests that gene targeting at the human HPRT locus proceeds with high fidelity.

At the HPRT locus in mouse ES cells we analyzed two targeted recombinants that arose from a replacement vector (Fig. 1*A*). Neither targeted recombinant contained a detectable mutation (data not shown).

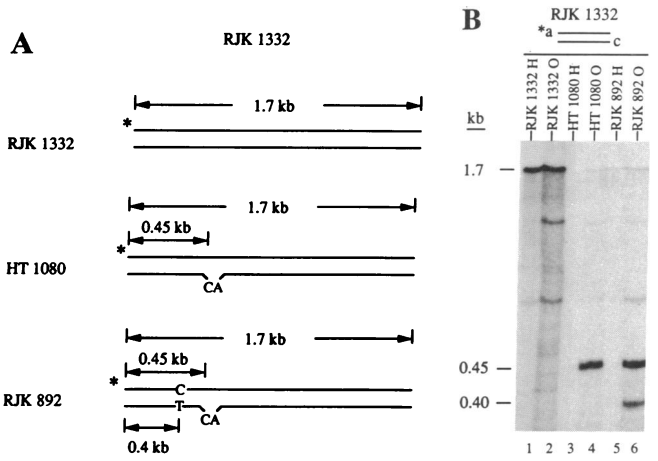

FIG. 2. Chemical mismatch cleavage reaction on control DNA. (*A*) Representation of heteroduplexes. RJK 1332 had a GT deletion and RJK 892 had a C → A point mutation in exon 3 (17). The probe was made by a PCR using radiolabeled primer a and nonlabeled primer c and RJK 1332 DNA as template. When the probe was annealed to the antisense HT1080 DNA, the wild-type sequence served as an insertional mutation at the site of deletion, and a product of 450 base pairs (bp) was expected to form if topological stress could lead to a cleavage at this site. When the same probe was annealed to another mutant, RJK 892, in addition to the cleavage product expected from HT1080 DNA, another product should be formed at 400 bp, corresponding to the C → A point mutation, i.e., C → T mismatch. Because the reaction is partial, the 1.7-kb wild-type product should be seen in all cases. (*B*) Demonstration of the cleavage products. The wild-type products at 1.7 kb existed in all lanes. There was no specific cleavage products when the probe was annealed to itself (lanes 1 and 2). However, expected cleavage products were generated when the probe was annealed to both HT1080 DNA (lane 4) and RJK 892 (lane 6). H, hydroxylamine; O, osmium tetroxide.

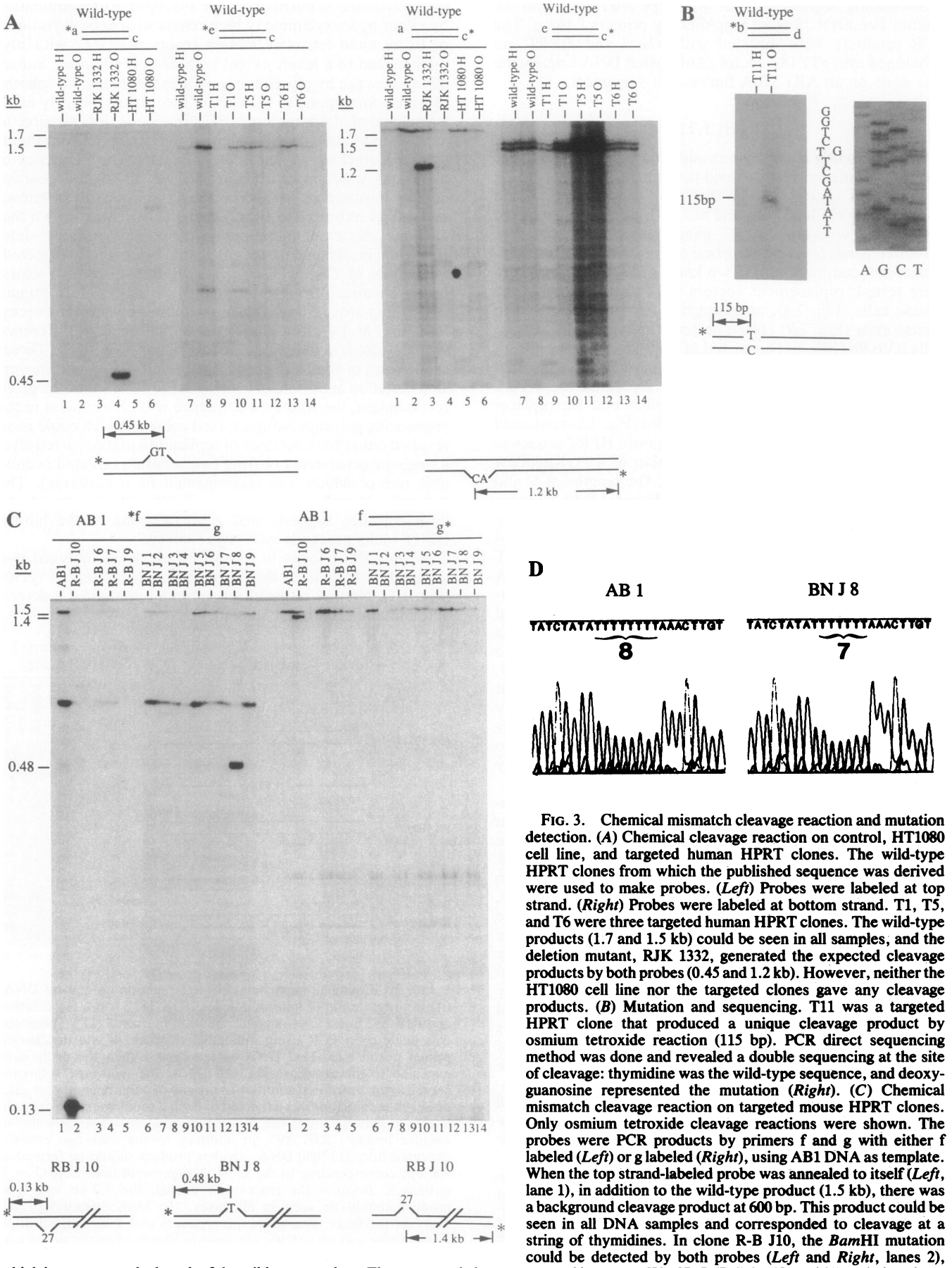

FIG. 3. Chemical mismatch cleavage reaction and mutation detection. (*A*) Chemical cleavage reaction on control, HT1080 cell line, and targeted human HPRT clones. The wild-type HPRT clones from which the published sequence was derived were used to make probes. (*Left*) Probes were labeled at top strand. (*Right*) Probes were labeled at bottom strand. T1, T5, and T6 were three targeted human HPRT clones. The wild-type products (1.7 and 1.5 kb) could be seen in all samples, and the deletion mutant, RJK 1332, generated the expected cleavage products by both probes (0.45 and 1.2 kb). However, neither the HT1080 cell line nor the targeted clones gave any cleavage products. (*B*) Mutation and sequencing. T11 was a targeted HPRT clone that produced a unique cleavage product by osmium tetroxide reaction (115 bp). PCR direct sequencing method was done and revealed a double sequencing at the site of cleavage: thymidine was the wild-type sequence, and deoxyguanosine represented the mutation (*Right*). (*C*) Chemical mismatch cleavage reaction on targeted mouse HPRT clones. Only osmium tetroxide cleavage reactions were shown. The probes were PCR products by primers f and g with either f labeled (*Left*) or g labeled (*Right*), using AB1 DNA as template. When the top strand-labeled probe was annealed to itself (*Left*, lane 1), in addition to the wild-type product (1.5 kb), there was a background cleavage product at 600 bp. This product could be seen in all DNA samples and corresponded to cleavage at a string of thymidines. In clone R-B J10, the *Bam*HI mutation could be detected by both probes (*Left* and *Right*, lanes 2), which in sum, gave the length of the wild-type product. Three targeted clones generated by vector IV6.8R-B (R-B J6, J7, and J9) and nine clones by vector IV6.8BN (BN J1–J9) were shown, only one cleavage product could be detected (*Left*, BN J8, lane 13). (*D*) Mutation sequencing. (*Left*) Wild-type sequence from AB1 cells. (*Right*) Mutant with a thymidine deletion in the eight-thymidine cluster.

Table 1. Summary of chemical mismatch cleavage results on targeted HPRT clones

Species	Human	Mouse
Cell line	HT 1080	AB 1
Class	Fibroblasts	ES cells
No. of clones		
Replacement	11	2
Insertion	0	31
Sequence analyzed on each clone (kb)	2.8	1.5
Total sequence analyzed (kb)	31	50
No. of mutations	1	1
Type	T → G transversion	Deletion of T

Analysis of Targeted Clones Generated with Insertion Vectors. The replacement vectors used in targeting the mouse and human HPRT genes were perfectly homologous to the chromosome. Because a high frequency of mutation was previously observed resulting from heteroduplex formation (6), we decided to examine the influence of limited heterology in our targeting vectors. We used targeted recombinants that arose from three different insertion vectors (Fig. 1*B*). One vector, IV6.8 was identical to the chromosomal sequence; a second, IV6.8R-B, had an internal *Eco*RI site changed to a *Bam*HI site by insertion of a linker (12); and the third, IV6.8BN, carried 13 nucleotides of heterology at each free end by linearizing with *Not* I. Ten products from each vector that generated the wild-type sequences in both duplications and one recombinant from IV6.8R-B that had the wild-type *Eco*RI site changed to the *Bam*HI site in both duplications (IV6.8R-B J10) were chosen for chemical cleavage analysis. These particular recombinants were selected because they were the most likely to have resulted from repair of heteroduplexes formed during the recombination event. All 31 recombinants were analyzed. One new mutation was detected in a recombinant derived from the vector IV6.8BN, which carried short terminal nonhomologies (Fig. 3*C*). Automated DNA sequencing of the PCR product showed deletion of a thymidine at the site of cleavage (Fig. 3*D*).

DISCUSSION

The chemical mismatch cleavage data derived from the analysis of all targeted clones are summarized in Table 1. Only two mutations were detected from a total of 80 kb around 44 targeted sites, indicating that targeted recombination, like intrachromosomal recombination (5), can be very accurate in mammalian cells. In mouse cells heterology at the free ends of the vector or at a single site within the vector did not have an effect on the frequency of mutation.

These results contrast with previous reports in which targeted recombination seemed mutagenic (6–8). For the results obtained by Thomas and Capecchi (6) and by Brinster *et al.* (8), the vectors were introduced by microinjection, and the vector and chromosomal sequences were not identical. Microinjection could conceivably disturb the normal fidelity of recombination, or, alternatively, the formation of mismatched heteroduplexes during recombination could induce an error-prone repair process. However, in our experiments, limited heterology either inside the homology or at the end of homology did not result in any increase in the frequency of mutation.

The mutations detected in our experiments could have arisen spontaneously in the cells before transfection or may have arisen during the recombination process or during electroporation of the DNA into the cells. While the source of our mutations is unclear, the low frequency observed in this study should not interfere with gene targeting as an approach to achieve accurate modifications of chromosomal loci in mammalian cells.

Gene targeting technology has moved beyond simple gene "knock-outs;" methods have now been developed to create very subtle modifications of mammalian genomes by homologous recombination (12, 18–20). The ability to create subtle mutations in ES cells and carry them into the mouse germ line is essential for detailed *in vivo* analysis of gene function and regulation and for creating animal models that precisely mimic mutations of human diseases. Here we studied the products of targeted recombination at the nucleotide sequence level and demonstrated that it is an extremely accurate process. Thus, precise modifications of mammalian genomes can be achieved by gene targeting with high fidelity. In addition, this work assures that targeted recombination can be used to correct a deficient gene in a predictable manner and, therefore, could be applied in human gene therapy.

We thank De Dieu, Jaime Rivera-Perez, and J. D. Wallace for expert technical assistance, and Al Edwards for providing human HPRT clones. We are grateful to Ann Davis and Xiangwei Wu for their helpful suggestions. This work was supported by grants from the National Institutes of Health (A.B. and J.H.W.), the Searle Scholars Program/Chicago Community trust (A.B.), the Cystic Fibrosis Foundation (A.B., J.H.W., and P.H.), the American Cancer Society (J.H.W.), and the March of Dimes Birth Defects Foundation (H.Z.).

1. McMahon, A. P. & Bradley, A. (1990) *Cell* **62,** 1073–1085.
2. Soriano, P., Montgomery, C., Geske, R. & Bradley, A. (1991) *Cell* **64,** 693–702.
3. Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. & Melton, D. W. (1989) *Cell* **56,** 313–321.
4. Zijlstra, M., Bix, M., Simister, N. E., Loring, J. M., Raulet, D. H. & Jaenisch, R. (1990) *Nature (London)* **344,** 742–746.
5. Stachelek, J. L. & Liskay, R. M. (1988) *Nucleic Acids Res.* **16,** 4069–4076.
6. Thomas, K. R. & Capecchi, M. R. (1986) *Nature (London)* **324,** 34–38.
7. Doetschman, T., Maeda, N. & Smithies, O. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 8583–8587.
8. Brinster, R. L., Braun, R. E., Lo, D., Avarbock, M. R., Oram, F. & Palmiter, R. D. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 7087–7091.
9. Rasheed, P. S., Nelson-Rees, W. A., Toth, E. M., Arnstein, P. & Gardner, M. B. (1974) *Cancer* **33,** 1027–1033.
10. Cotton, R. G. H., Rodrigues, N. R. & Campbell, R. D. (1988) *Proc. Natl. Acad. Sci. USA* **85,** 4397–4401.
11. Grompe, M., Muzny, D. M. & Caskey, C. T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5888–5892.
12. Hasty, E. P., Ramirez-Solis, R., Krumlauf, R. & Bradley, A. (1991) *Nature (London)* **350,** 243–246.
13. Gibbs, R. A., Nguyen, P. N. & Caskey, C. T. (1991) in *Methods in Molecular Biology*, ed. Mathews, C. (Humana Press, London), in press.
14. Thomas, K. R. & Capecchi, M. R. (1987) *Cell* **51,** 503–512.
15. Edwards, A., Voss, H., Rice, P., Civitello, A., Stegemann, J., Schwager, C., Zimmermann, J., Erfle, H., Caskey, C. T. & Ansorge, W. (1990) *Genomics* **6,** 593–608.
16. Melton, D. W., Konecki, D. S., Brennand, J. & Caskey, C. T. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2147–2151.
17. Gibbs, R. A., Nguyen, P. N., McBride, C. J., Koepf, S. M. & Caskey, C. T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 1919–1923.
18. Zimmer, A. & Gruss, P. (1989) *Nature (London)* **338,** 150–153.
19. Steeg, C. M., Ellis, J. & Bernstein, A. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 4680–4684.
20. Valancius, V. & Smithies, O. (1991) *Mol. Cell. Biol.* **11,** 1402–1408.