## Targeted correction of a major histocompatibility class II  $E_{\alpha}$  gene by DNA microinjected into mouse eggs

(transgenic mice/homologous recombination/gene repair/mutation)

RALPH L. BRINSTER\*<sup>†</sup>, ROBERT E. BRAUN<sup>‡</sup>, DAVID LO\*, MARY R. AVARBOCK\*, FELICITY ORAM<sup>‡</sup>, AND RICHARD D. PALMITERI

\*Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, PA 19104; and tHoward Hughes Medical Institute, Department of Biochemistry, University of Washington, Seattle, WA <sup>98195</sup>

Contributed by Ralph L. Brinster, June 22, 1989

ABSTRACT DNA molecules containing the <sup>5</sup>' end of <sup>a</sup> functional major histocompatibility class II  $E_{\alpha}$  gene were injected into mouse eggs bearing  $E_{\alpha}$  genes with 630-base-pair (bp) deletions encompassing the promoter and first exon. The deletion was corrected by homologous recombination in <sup>1</sup> of about 500 transgenic mice that incorporated the injected DNA. The corrected  $E_{\alpha}$  gene was transmitted to progeny, which were bred to homozygosity. Southern blot analysis, polymerase chain reaction amplification of the DNA spanning the deletion, and sequence analysis revealed that the corrected allele resembles the wild-type  $E_{\alpha}$  gene. At sites of single-base-pair polymorphisms, there was apparently random conversion to either the donor or recipient sequence. In addition, many point mutations were introduced. mRNAs were produced from the corrected allele in a tissue-specific manner, but their sizes were different from the wild-type allele, and they did not produce detectable  $E_{\alpha}$  protein. This experiment demonstrates the feasibility of targeting foreign DNA to <sup>a</sup> gene that is completely inactive in fertilized mouse eggs.

Targeting foreign DNA to specific chromosomal sites by homologous recombination provides an invaluable genetic tool for studying gene function and correcting genetic defects. In some organisms (S. cerevisiae, Neurospora, and Dictyostelium), homologous recombination is common, whereas in others (Drosophila and mouse), nonhomologous integration is the rule (1). In those organisms where nonhomologous integration predominates, achieving targeted insertion requires effective selection or screening strategies or a method for reducing the nonhomologous events.

Although DNA introduced into mammalian somatic cells (2-4) or injected into mouse eggs (5) recombines with coinjected molecules quite readily, the foreign DNA typically integrates randomly within the chromosomes (1-5). Several clever approaches have been developed to study the parameters that influence homologous recombination within specific chromosomal loci (6-10). In many of these studies, a selectable gene was used to facilitate the analysis. With these approaches there is about 1 homologous recombination event for every  $10<sup>3</sup>$  nonhomologous integration events  $(1, 11)$ .

There are some applications for homologous recombination in tissue culture cells, but for many experiments, the ultimate goal is to alter both alleles and assess the effects in the developing organism. One approach has been to take advantage of embryonal carcinoma or embryonal stem (ES) cells that can be manipulated in culture by transfection or retroviral infection, selected for desired phenotype, and then introduced into the blastocyst where they can contribute to the formation of the developing mouse (11-15). Colonization

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.

of the germ line of the resulting chimeric mice is more efficient when ES cells are used, providing the possibility of breeding the offspring that carry a targeted allele to homozygosity and studying the effects (11, 15). Methods have been developed either to select for targeting events to genes that are not naturally selectable or to fractionate cell populations and identify desired recombinants by the polymerase chain reaction (PCR) technique. Using these approaches, the Hprt, Int-2, En-2, and Hox-1.1 genes have been disrupted in ES cells, and chimeric mice have been produced (16-18). However, incorporation of ES cells from this type of experiment into the germ line was only demonstrated in one experiment in which a defective Hprt gene was corrected in a newly derived ES cell line (19). Conceivably, using fresh cells and a selectable gene minimized the opportunity for chromosomal aberrations that are believed to interfere with germ-line contribution of ES cells.

The frequency of homologous integration into various loci<br>has ranged from  $10^{-2}$  to  $10^{-5}$  when DNA was either injected or electroporated into ES cells (12-14, 16-19). With a frequency of 1% or better, direct microinjection of DNA into fertilized mouse eggs would represent a feasible alternative to the ES cell approach. The major advantage would be that transgenes introduced by microinjection into eggs are usually represented in the germ line. Thus, we set out to ascertain the feasibility of this direct approach.

The major histocompatibility (MHC) class II  $E_{\alpha}$  gene is mutated in several strains of mice, including the C57BL/6 and SJL strains that we routinely use to make transgenic mice (20). The  $E^b_\alpha$  (C57BL/6) and  $E^s_\alpha$  (SJL) alleles each bear a 630-bp deletion that removes the first exon and part of the promoter (21). The normal, functional  $E^d_{\alpha}$  allele (from BALB/ c mice) has been cloned on a cosmid (21), and injection of this cosmid or smaller subclones into mouse eggs has been shown to completely correct the genetic defect in mutant mice by complementation (22-24). We chose to extend these studies by attempting to correct this genetic defect by homologous recombination using only part of the gene spanning the deletion.

## MATERIALS AND METHODS

Animals and DNA Microinjection. Eggs from  $F_1$  (C57BL/6  $\times$  SJL) females were fertilized by F<sub>1</sub> (C57BL/6  $\times$  SJL) males and injected with the DNA preparations shown in Fig. <sup>1</sup> as described (20). Transgenic mice were identified by dot hybridization using probe B.

DNA Analysis. DNA was prepared from tails or other tissues by the sodium dodecyl sulfate/proteinase K method as described (20). For Southern blots, 5  $\mu$ g of DNA was

Abbreviations: ES cells, embryonal stem cells; PCR, polymerase chain reaction.

tTo whom reprint requests should be addressed.



FIG. 1. Restriction map of the MHC class II  $E_{\alpha}^{d}$  gene showing the location of microinjected DNA fragments, nick-translated probes, and PCR primers. The DNA fragments that were microinjected are designated by numbers: 454 and 620 are the same Sal <sup>I</sup> fragment except that 620 was isolated from a subclone of this fragment in pUC, whereas 454 was isolated from a  $Bgl$  I-HindIII plasmid containing the entire  $E_{\alpha}^{d}$  gene; 647 was isolated from a subclone of the 8.9-kilobase (kb)  $EcoRI$  fragment in pUC; and 588 was prepared by ligating the  $Bgl$ <sup>I</sup> site to EcoRI (shown as a dotted line) and then cutting with Pst <sup>I</sup> (it was propagated by cloning it into the  $Pst$  I site of pUC). The probes were each subcloned into  $pUC$ : A is a Sst I-Bgl I fragment, B is an Acc I-Sst I fragment, and C is an EcoRI-Nsi I fragment. The approximate location of PCR primers is indicated; more details are provided in Materials and Methods. Brackets show the location of the promoter/exon <sup>1</sup> deletion in C57BL/6 and SJL mice.

digested overnight with 10-15 units of restriction enzyme, electrophoresed on a 0.8% agarose gel, transferred to nitrocellulose, and hybridized with nick-translated probes by standard procedures (20, 25). For PCR, 2  $\mu$ g of DNA sample was mixed with reagents recommended in the Cetus protocol along with pairs of primers indicated in Fig. 1:

- $d = (-568)$  5'-TGTCTAGCCCACTGCAAAAGG,
- $e = (-54)$  5'-ATTCTGCCTCAGTCTGCGAT,  $f = (+922)$  5'-CAGACTTGCTCTTGTGTGAC
- $g = (+2678)$  5'-TAGAAGGGCAAGCAAGACCTC;

the numbers in parentheses indicate the position of the <sup>5</sup>' end of the oligonucleotide on the  $E_{\alpha}^{d}$  sequence, where the adenosine of the initiation codon represents position  $+1$ . DNA

was denatured at 93°C for <sup>1</sup> min, annealed at 48°C for 20 sec, and elongated at 72°C for 2 or 3 min for 30 cycles; then a sample was electrophoresed on a 0.8% agarose gel and stained with ethidium bromide.

Cloning and Sequencing the PCR Amplification Products. Aliquots of the PCR reaction were phosphorylated with phage T4 polynucleotide kinase, digested with Sst I, and cloned into a Sma I/Sst I-cut pUC18 plasmid. Plasmids with either the <sup>5</sup>' or <sup>3</sup>' fragments were identified by restriction analysis; CsCl preparations were prepared, and they were used for the dideoxy chain-termination method that used Sequenase with double-stranded DNA templates and either the universal pUC18 primer or primers complementary to the  $E_{\alpha}^{d}$  sequence.

RNA Analysis. Total RNA was prepared by the LiCl/ guanidine thiocyanate method (26); aliquots (3 or 15  $\mu$ g) were electrophoresed on 1.0% formaldehyde/agarose gels, and the RNA was transferred to nitrocellulose and hybridized with nick-translated probes prepared as for Southern hybridization.

## RESULTS

Detection of Homologous Recombination at the  $E_{\alpha}$  Locus. An intact  $E_a$  gene, cloned from BALB/c mice (d haplotype) by Hyldig-Nielsen et al. (21), was used as <sup>a</sup> source of DNA for targeted correction of the  $E_{\alpha}$  gene in embryos resulting from microinjected fertilized eggs. Both replacement- and insertion-type DNA molecules (see ref. <sup>15</sup> for description) were tested for their ability to correct the  $E_{\alpha}$  deletion. Because we were trying to correct the expression of a

promotorless gene and because plasmid sequences sometimes inhibit expression, the injected DNA was excised from the bacterial vector prior to microinjection. The three DNA molecules that were tested are shown below a map of the  $E_{\alpha}$ gene in Fig. 1. DNAs 454, 620, and 647 are replacement-type molecules, whereas DNA <sup>588</sup> is an insertion-type molecule.

Each of the DNAs was diluted to about 2 ng/ $\mu$ l, and about 2 pl was injected into the pronuclei of  $F_2$  mouse eggs; this corresponds to an injection of 400-1400 molecules of DNA. Transgenic mice that had incorporated one or more copies of the injected DNA were identified by dot hybridization of tail DNA by using nick-translated probe B (see Fig. 1) that lies within the 630-base-pair (bp) deletion of the parental strains. The frequency of obtaining transgenic mice with these DNAs was normal; a total of 506 transgenic mice were identified among the 1841 mice that were born, for a frequency of  $27.5\%$ (Table 1). These mice were obtained by injecting about 10,000 eggs over <sup>a</sup> period of <sup>14</sup> months. DNA from each of these transgenic mice was analyzed by Southern blot for evidence of homologous recombination. Most of the nontransgenic tail DNA samples were also analyzed as pools of about <sup>10</sup> DNAs by PCR with primers <sup>e</sup> and <sup>f</sup> or <sup>e</sup> and g (see Fig. 1) to avoid missing any low copy or mosaic founders.

For Southern analysis, a probe lying outside of the region spanned by the plasmid from which the injected DNA was isolated was used. For DNAs <sup>454</sup> and 620, probe A was used with tail genomic DNAs restricted with Sst I; whereas with DNAs <sup>588</sup> and 647, probe C was used with genomic DNAs restricted with Nsi <sup>I</sup> (see Fig. 1). Thus, the Southern procedure examines the endogenous  $E_{\alpha}$  locus and should reveal only the  $E_{\alpha}$  s and b alleles unless the injected DNA has recombined into the  $E_{\alpha}$  locus. New bands corresponding in size to the  $E^d_{\alpha}$ -positive control would suggest correct targeting, while any other bands might signify some other change within the locus. Only one targeting event, resulting from the <sup>454</sup> DNA preparation, was detected. The data below establish that it was a bona fide targeting event.

Mouse <sup>944</sup> Has <sup>a</sup> Targeted Gene Correction. When the DNA from mouse 944-1 was digested with Sst <sup>I</sup> and analyzed with probe A, a band identical to the d allele was observed along with a band corresponding to  $b$  or  $s$  alleles (data not shown). This founder female was first bred to a C57BL/6 male and then to an  $F_1$  (C57BL/6  $\times$  SJL) male. Eight of the first 26 offspring inherited the new allele. Two pairs of offspring with the new allele were bred to generate homozygous mice that were identified with probe  $\overline{A}$  on Southern blots of genomic DNA digested with Sst I. When these blots were hybridized with the original <sup>454</sup> DNA fragment that was microinjected, only the two Sst I bands corresponding to the  $E_{\alpha}$  locus hybridized, indicating that there was only a single insert of foreign DNA into the genome of this line of mice. We considered the possibility that the correction event might be due to gene conversion from the  $A_{\alpha}$  locus; however, the sequence of the  $A_{\alpha}$  gene is substantially different from that of  $E_{\alpha}$  and fails to cross hybridize with probes A, B, or C under the standard conditions used for these Southern blots.

Table 1. Targeted correction of an  $E_{\alpha}$  gene in transgenic mice

DNA*	Gene size, kb	Eggs injected, no.	<b>Pups</b> born. no.	Mice	
				Transgenic, no.	Targeted, <sup>†</sup> no.
454	2.56	2.049	341	84	
620	2.56	1.699	366	94	0
588	4.6	2,420	459	84	0
647	8.9	4.434	675	244	0
<b>Totals</b>		10,602	1,841	506	

\*See Fig. 1.

tNumber of transgenic mice with a targeted integration.

The previous results suggest that only a single copy of  $E^d_{\alpha}$ was inserted. To address this question directly, DNA from <sup>a</sup> hemizygous offspring (mouse 944-1-8) and two heterozygous controls (one carrying  $b$  and  $s$  alleles, the other carrying  $d$  and s alleles) was digested with HindIII, HindIII/BamHI, or Pvu II, enzymes that do not cut within the injected DNA 454. In each case, the 944-1-8 heterozygote gave two bands, one the size of the  $s$  allele and one the size of the  $d$  allele, which is about 600 bp larger (Fig. 2). These results are consistent with the correction of either the  $E_a$  b or s allele without any other insertion of foreign DNA.

Because the founder 944 mouse was the result of mating  $F_1$ hybrid parents, either the  $E_{\alpha}$  b or s allele could have been corrected. To ascertain which allele was corrected, we first identified a Pst I polymorphism between the  $E_{\alpha}$  s and b alleles by using probe C and then examined a mouse homozygous for the 944 allele. This experiment revealed that the s allele was corrected (data not shown).

To characterize the targeting event in more detail, we used PCR to amplify DNA from the  $E_{\alpha}$  locus of line 944 heterozygotes and homozygotes as well as SJL and BALB/c  $(E_{\alpha}^{d})$ controls using primers d and f or e and g (see Fig. 1). The sizes of the amplification products from the line 944 mice agreed with that of the  $E^d_{\alpha}$ -positive control with both sets of primers (data not shown). The DNAs from one control SJL mouse (homozygous for the s allele) and three different 944 homozygotes were amplified with primers d and f and cloned into pUC18 vector. Clones containing  $E_{\alpha}$  inserts were identified and sequenced. In addition, part of the plasmid preparation used to isolate the <sup>454</sup> DNA for injection was sequenced by using primers d and f.

Sequence analysis revealed the following seven facts: (i) The 600 bp of sequence obtained with primers d and f from the  $E_{\alpha}^{d}$  plasmid was identical to that published by Hyldig-Nielson *et al.* (21). (*ii*) The 630-bp deletion ( $-398$  to  $+232$ ) in the s allele is the same as that in the  $E_a$  b allele (27). (iii) The boundaries of the  $E_{\alpha}^{s}$  deletion were precisely corrected in the 944 line. (iv) There are three polymorphic differences between  $E_{\alpha}^{s}$  and  $E_{\alpha}^{d}$  alleles in the 149 bp 5<sup>7</sup> of the deletion and nine polymorphisms in the 670 bp <sup>3</sup>' of the deletion as shown schematically in Fig. 3.  $(v)$  Comparison of the sequence with that of the  $b$  allele confirms that the  $s$  allele was corrected



FIG. 2. Southern blot of DNA from <sup>a</sup> mouse homozygous for the deletion (b-s, an  $F_1$  hybrid of C57BL/6  $\times$  SJL), a control hemizygous for the deletion (d-s, an  $F_1$  hybrid of BALB/c  $\times$  SJL), and an experimental mouse hemizygous for the injected DNA (944-1-8). The diagram shows the location of relevant restriction sites and probe A (see Fig. 1 for more information); the location of the deletion is shown by brackets. P, Pvu II; B, BamHI; H, HindIII.

rather than the  $b$  allele (data not shown). (vi) At 5 of the 12 sites that are polymorphic between  $E_{\alpha}^{d}$  and  $E_{\alpha}^{s}$ , the sequence matches that of the d allele, and at 6 of 12 sites the sequence matches that of the s allele; these polymorphisms appear to be distributed randomly in the line 944 allele, suggesting that a heteroduplex was formed between these two molecules, and information from either strand was used to correct the sequence during DNA repair. (vii) Fifteen new mutations (labeled <sup>n</sup> or N in Fig. 3) were introduced into the <sup>944</sup> allele in the 1491 bp that were sequenced; one of these (N) maps at a site of  $s/d$  polymorphism.

Gene Expression from the Corrected Line 944 Ailele. Based on the DNA analysis, we expected that transcription might be restored to the line 944 allele because the promoter and first exon were restored. We measured  $E_{\alpha}$  RNA abundance in RNA isolated from the spleen and thymus of the line 944 mice by solution hybridization with oligonucleotides complementary to exons 2, 3, and 4 of  $E^d_\alpha$  mRNA. These data indicated that the amount of <sup>944</sup> allele mRNA was about 5% of that of an  $E^d_{\alpha}$  allele in BALB/c mice (data not shown). Northern blot analysis revealed two prominent bands of approximately equal intensity in spleen or thymus RNA from positive control mice (Fig. 4); the lower one corresponds in size to that reported by others, whereas the higher one corresponds to a weak band previously described (24). The spleen and thymus RNAs from three different homozygous line 944 mice reveal three major hybridizing bands: one is the same size as the larger band, one is slightly larger than the lower band, a third band lies in between them, and there are several minor bands. The amount of hybridization to these bands was always much less than to the BALB/c (d allele) control when the mice were treated similarly (data not shown). In the experiment shown in Fig. 4, the line 944 mice were stimulated with Freund's complete adjuvant, which induces class II gene expression, whereas the controls were not, which may account for the variable increase in  $E_{\alpha}$  mRNA in the line 944 samples. No expression of the 944 allele was detected in liver, kidney, or heart, indicating that the appropriate tissue specificity was maintained (Fig. 4).

To ascertain whether the line 944  $E_{\alpha}$  mRNA could code for functional protein, thereby restoring I-E expression on antigen-presenting cells, sections of spleen from homozygous mice either treated with Freund's complete adjuvant or untreated were processed for immunocytochemistry by using <sup>a</sup> monoclonal antibody specific for I-E molecules. No staining was observed, whereas the spleens from mice bearing the d allele gave a clear signal. For a more sensitive assay, line <sup>944</sup> mice were tested for tolerance to the I-E molecule. We have previously shown that expression of the I-E molecule in transgenic mice readily induces tolerance (29). T lymphocytes from line 944 mice generated strong proliferative responses to I-E-positive stimulator cells (not shown). The responses were equivalent to those from control I-E-negative mice, indicating that no tolerance was induced to I-E. These results suggest that the RNA transcripts from the <sup>944</sup> allele were unable to be translated into the functional  $E_{\alpha}$  protein.

## DISCUSSION

We have demonstrated that it is feasible to target foreign DNA to <sup>a</sup> specific locus of the mouse genome in fertilized mouse eggs. Because there has only been one targeting event out of more than 1800 mice born, about 500 of which were transgenic, we cannot estimate the frequency of homologous recombination. Nevertheless, the fact that we obtained one is a significant finding because there are at least two aspects of our experimental design that made it a formidable endeavor. First, the gene we were trying to target is completely inactive because it lacks a promoter. Mansour et al. (16) have provided evidence that targeting the active Hprt gene in ES



FIG. 3. Summary of DNA sequence analysis of d, s, and line 944  $E_{\alpha}$  alleles. DNA from an SJL mouse (s allele) and three homozygous line 944 mice were separately subjected to PCR [with primers d and <sup>f</sup> (see Fig. 1)]. The products were phosphorylated by using T4 kinase, digested with Sst I, and cloned between Sst I and Sma I sites of pUC for DNA sequence analysis; three independent clones representing the 5' and 3' fragments were sequenced by the dideoxy method as applied to double-stranded DNA. The top line shows the  $E_{\alpha}$  gene with the bracket indicating the deletion s and b alleles; the numbering system uses the A of the initiation codon as  $+1$ . The second line shows the injected Sal I fragment. The next line shows the location (with coordinates) of all of the polymorphic sites when comparing the sequences of the  $d$ ,  $s$ , and line 944 alleles; above the line the sites that are polymorphic between s and d alleles are noted with these same letters; below the line new mutations are indicated by n where the sequence of line 944 does not match either the s and d alleles; N indicates one polymorphic site that also was mutated. The nucleotides found at each of these positions are indicated in the table at the bottom (- indicates a deletion). The sequence of each of the three clones derived from SJL mice or line 944 homozygotes were identical with one exception, which was probably a Taq polymerase error.

cells is about 10-fold more efficient than targeting the relatively inactive Int-2 gene. However, targeting DNA to the  $\beta$ -globin locus was observed with similar frequency regardless of whether the cells were expressing the globin gene (7). Second, Waldman and Liskay (30) have shown that efficient recombination in mammalian cells appears to require be-



FIG. 4. Hybridization-blot analysis of RNA from mice homozygous for the line 944, b, s, or  $dE_{\alpha}$  alleles. Total RNA (3 or 15  $\mu$ g) from various tissues (ht, heart; ki, kidney; li, liver; sp, spleen; th, thymus) was electrophoresed on a denaturing 1.5% agarose gel, transferred to nitrocellulose, and hybridized with a 2.5-kb EcoRI-HindIII probe that carries exons 3-5. RNA was isolated from three different mice homozygous for the line 944 allele <sup>1</sup> week after stimulation with Freund's complete adjuvant (28).

intrachromosomal recombination test. In the 1500 bp that we sequenced, there were no regions of identity as large as this critical distance; hence, it is likely that these polymorphisms lowered the targeting efficiency.

When we analyzed the transgenic mice derived from injected DNA <sup>454</sup> with probe C, we observed that about 20% had an altered  $E_{\alpha}$  allele; however, this turned out to be an artifact due to contamination of the DNA <sup>454</sup> Sal <sup>I</sup> fragment with DNA from the plasmid from which it was isolated. Although this contamination was not visible on a gel, dot hybridization indicated that it represented about 1% of the DNA <sup>454</sup> preparation. Because some transgenic mice have hundreds of copies of the transgene, the possibility of incorporating contaminating DNA along with the DNA <sup>454</sup> fragment was high and occasionally that DNA would hybridize to probe C, generating a new band. When we realized the source of this artifact, we subcloned the DNA <sup>454</sup> to generate DNA 620. Because the <sup>944</sup> line of mice was derived from DNA 454, it is possible that other DNA from the plasmid may have participated in the recombination; nevertheless, all of our analysis indicates that it was a bona fide targeting event.

The event that gave rise to the 944 mouse line is remarkable in several regards. It appears to be the result of either a double recombination or a gene conversion event in which the 630-bp deletion and an uncertain amount of flanking DNA sequence was corrected. We have found no evidence of any other foreign DNA inserted into the genome of this line of mice. Thus, unlike the usual situation in producing transgenic mice, <sup>a</sup> large head-to-tail array of injected DNA was not

integrated into the mouse chromosome. The size of the correction patch was not determined, but it clearly extends throughout the 1500-bp region that we sequenced. At the outset, we imagined that there would be a contiguous patch of d allele sequence embedded in the <sup>s</sup> allele chromosome; however, the sequence of the <sup>944</sup> DNA indicates that the polymorphic sites were corrected to match either  $d$  or  $s$  in an apparently random manner right up to the boundaries of the deletion. One explanation for this observation is that a heteroduplex between one strand of s allele DNA and one strand of d allele DNA formed as an intermediate in the recombination event and that error-prone repair enzymes had no strand preference in correcting the heteroduplex. The lack of strand preference is surprising because the incoming DNA would have contained bacterial DNA methylation, whereas the target DNA probably had 5-methylcytosine modifications; thus, the strands could have been distinguished on this basis.

In addition to the patchy correction of  $d$  and  $s$  polymorphisms, there was an unexpectedly large number of new mutations. We identified <sup>15</sup> new mutations in the <sup>1500</sup> bp that were sequenced. These mutations are not due to PCR amplification with Thermus aquaticus (Taq) polymerase because we compared the sequence from three independent line <sup>944</sup> mice and they were identical in all but one case. We cannot rule out the possibility that these mutations occurred prior to the recombination event because few sequences have been published comparing DNA that was microinjected with that recovered from transgenic mice. In one case, the sequence of a transgene bearing a metallothionein-thymidine kinase gene in pBR322 was recovered and sequenced; it was not mutated in the regions that were sequenced (25). Thus, it is likely that the mutagenesis occurred during the recombination event. The mutagenesis may have extended beyond the boundaries of the 454 fragment because we have noted a loss of the Nsi I site in the third intron (see Fig. 1). The d-like polymorphism at position +866 (Fig. 3) might also be explained by mutation or by contamination with flanking DNA. New mutations have been detected previously in recombination experiments involving mammalian cells (8). Analysis of several more recombinants will be required to ascertain whether this is a general phenomenon.

The corrected allele in the 944 line is expressed at a low level compared to the d allele. At the RNA level, we estimated that expression is 5-10% of normal in spleen and thymus, but it has the same cell-type specificity as the  $d$ allele. The appropriate cell specificity argues that the responsible cis-acting elements are intact (31, 32), whereas the low level of expression could be attributed to mutations in essential promoter elements. We have compared the sequence of the allele in line 944 with that of the  $X$  and  $Y$  boxes that have been implicated as important promoter elements (33) and they are intact; however, there may be other important elements that are mutated. Alternatively, the inactive <sup>s</sup> allele may have become modified (e.g., methylated) in a way that reduces expression, and some of these modifications may have been retained after the targeting event and continue to suppress expression. We also have noted that the size of the transcripts from the line 944 allele are altered; this might reflect changes in sequences that affect RNA splicing or polyadenylylation. Sequence analysis of the entire line 944 allele might provide insight into these possiblities. Inappropriate RNA processing or mutations in the  $E_{\alpha}$  coding region may account for the lack of evidence for  $E_\alpha$  protein synthesis.

If the high mutation frequency is a common event in mammalian cells, it will mitigate the usefulness of homologous recombination for gene correction. If this technique is ever applied to gene therapy, it will be essential to document

not only perfect correction at the DNA level but also the absence of any other deleterious changes in the locus prior to therapeutic transfer of cells. In contrast, for gene disruption experiments, a high mutation frequency probably would not present a serious problem.

If procedures can be found that increase the frequency of homologous recombination to 1% or better, then direct microinjection of DNA into fertilized eggs could become <sup>a</sup> valuable method for targeting specific genes. This approach has the advantage that there is little uncertainty that the targeted gene will be carried in the germ line.

We thank R. Flavell and L. Burkly for the  $E_{\alpha}$  clones and helpful discussions and also J. Papamatheakis for sequence information on the  $E_{\alpha}^{b}$  allele. This work was supported in part by Grants HD-23657 and HD-09172 from the National Institutes of Health.

- Roth, D. & Wilson, J. (1988) Genetic Recombination, eds. Kucherlapati, R. & Smith, G. R. (Am. Soc. Microbiol., Washington, DC), pp. 621-653.
- 2. Folger, K. R., Wong, E. A., Wahl, G. & Capecchi, M. R. (1982) Mol. Cell. Biol. 2, 1372-1387.
- 3. Subramani, S. & Berg, P. (1983) Mol. Cell. Biol. 3, 1040-1052.<br>4. Smithies, O., Koralewski, M. A., Song, K.-Y. & Kucherlana
- Smithies, O., Koralewski, M. A., Song, K.-Y. & Kucherlapati, R. S. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 161-170.
- 5. Palmiter, R. D., Hammer, R. E. & Brinster, R. L. (1985) Genetic Manipulation of the Early Mammalian Embryo, eds. Costantini, F. & Jaenisch, R. (Cold Spring Harbor Lab., Cold Spring Harbor, NY), pp. 123-132.
- 6. Folger, K., Thomas, K. & Capecchi, M. R. (1984) Cold Spring Harbor Symp. Quant. Biol. 49, 123-138.
- 7. Smithies, O., Gregg, R. G., Boggs, S. S., Koralewski, M. A. & Kucherlapati, R. S. (1985) Nature (London) 317, 230-234.
- 8. Thomas, K. R. & Capecchi, M. R. (1986) Nature (London) 324, 34-38.<br>9. Thomas, K. R., Folger, K. R. & Capecchi, M. R. (1986) Cell 44, 419-
- 9. Thomas, K. R., Folger, K. R. & Capecchi, M. R. (1986) Cell 44, 419- 428.
- 10. Jasin, M. & Berg, P. (1988) Genes Dev. 2, 1353-1363.<br>11. Frohman, M. A. & Martin, G. R. (1989) Cell 56, 145-
- 11. Frohman, M. A. & Martin, G. R. (1989) Cell 56, 145-147.<br>12. Thomas, K. R. & Capecchi, M. R. (1987) Cell 51, 503-512
- 12. Thomas, K. R. & Capecchi, M. R. (1987) Cell 51, 503-512.<br>13. Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M.
- Doetschman, T., Gregg, R. G., Maeda, N., Hooper, M. L., Melton, D. W., Thompson, S. & Smithies, 0. (1987) Nature (London) 330, 576-578.
- 14. Doetschman, T., Maeda, N. & Smithies, 0. (1988) Proc. NatI. Acad. Sci. USA 85, 8583-8587.
- 15. Capecchi, M. R. (1989) Trends Genet. 5, 70-76.
- 16. Mansour, S. L., Thomas, K. R. & Capecchi, M. R. (1988) Nature (London) 336, 348-352.
- 17. Joyner, A. L., Skarnes, W. A. & Rossant, J. (1989) Nature (London) 338, 153-156.
- 18. Zimmer, A. & Gruss, P. (1989) Nature (London) 338, 150-153.<br>19. Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. &
- Thompson, S., Clarke, A. R., Pow, A. M., Hooper, M. L. & Melton, D. W. (1989) Cell 56, 313-321.
- 20. Brinster, R. L., Chen, H. Y., Trumbauer, M. E., Yagle, M. K. & Palmiter, R. D. (1985) Proc. Natl. Acad. Sci. USA 82, 4438-4442.
- 21. Hyldig-Nielsen, J. J., Schenning, L., Hammerling, U., Widmark, E., Heldin, E., Lind, P., Serveius, B., Lund, T., Flavell, R., Lee, J. S., Trowsdale, J., Schreier, P. S., Zablizky, F., Larhammar, D., Peterson, P. A. & Rask, L. (1983) Nucleic Acids Res. 11, 5055-5071.
- 22. LeMeur, M., Gerlinger, P., Benoist, C. & Mathis, D. (1985) Nature (London) 316, 38-42.
- 23. Yamamura, K., Kikutani, H., Folsom, V., Clayton, L., Kimoto, M., Akira, S., Kashiwamura, S., Tonegawa, S. & Kishimoto, T. (1985) Nature (London) 316, 67-69.
- 24. Pinkert, C. A., Widera, G., Cowing, C., Heber-Katz, E., Palmiter, R. D., Flavell, R. & Brinster, R. L. (1985) EMBO J. 4, 2225-2230.
- 
- 25. Wilkie, T. M. & Palmiter, R. D. (1987) Mol. Cell. Biol. 7, 1646-1655.<br>26. Cathala, G., Savouret, J. F., Mendez, B. L., Karin, M., Martial, J. A. Cathala, G., Savouret, J. F., Mendez, B. L., Karin, M., Martial, J. A. & Baxter, J. D. (1983) DNA 2, 329-335.
- 27. Dembič, D., Ayane, M., Klein, J., Steinmetz, M., Benoist, C.O. & Mathis, D. J. (1985) *EMBO J.* 4, 127-131.
- 28. Corradin, G., Etlinger, H. M. & Chiller, J. M. (1977) J. Immunol. 119, 1048-1053.
- 29. Widera, G., Burkly, L. C., Pinkert, C. A., Bottger, E., Cowing, C., Palmiter, R. D., Brinster, R. L. & Flavell, R. A. (1987) Cell 51, 175-187.
- 30. Waldman, A. S. & Liskay, R. M. (1988) Mol. Cell. Biol. 8, 5350–5357.<br>31. Dorn. A. Fehling. H. J. Koch. W. Le Meur. M. Gerlinger. P., Benoist. Dorn, A., Fehling, H. J., Koch, W., Le Meur, M., Gerlinger, P., Benoist,
- C. & Mathis, D. (1988) Mol. Cell. Biol. 8, 3975-3987.
- 32. Burkly, L. C., Lo, D., Cowing, C., Palmiter, R. D., Brinster, R. L. & Flavell, R. A. (1989) J. Immunol. 142, 2081-2088.
- 33. Dorn, A., Durand, B., Marfing, C., Le Meur, M., Benoist, C. & Mathis, D. (1987) Proc. Natl. Acad. Sci. USA 84, 6249-6253.