MOLECULAR AND CELLULAR BIOLOGY, Apr. 1993, p. 2134–2140
0270-7306/93/042134-07$02.00/0

Vol. 13, No. 4

Location of Crossovers during Gene Targeting with Insertion and Replacement Vectors

CHUXIA DENG, KIRK R. THOMAS, AND MARIO R. CAPECCHI*

Howard Hughes Medical Institute, Eccles Institute of Human Genetics, University of Utah School of Medicine, Salt Lake City, Utah 84112

Received 8 December 1992/Returned for modification 21 January 1993/Accepted 25 January 1993

Gene targeting was used to introduce nonselectable genetic changes into chromosomal loci in mouse embryo-derived stem cells. The nonselectable markers were linked to a selectable marker in both insertion- and replacement-type vectors, and the transfer of the two elements to the *Hprt* locus was assayed. When insertion vectors were used as substrates, the frequency of transfer was highly dependent upon the distance between the nonselectable marker and the double-strand break in the vector. A marker located close to the vector ends was frequently lost, suggesting that a double-strand gap repair activity is involved in vector integration. When replacement vectors were used, cotransfer of a selectable marker and a nonselectable marker 3 kb apart was over 50%, suggesting that recombination between vector and target often occurs near the ends of the vector. To illustrate the use of replacement vectors to transfer specific mutations to the genome, we describe targeting of the *ΔF508* mutation to the *CFTR* gene in mouse embryo-derived stem cells.

Gene targeting in mouse embryo-derived stem (ES) cells provides the means for generating mice of virtually any desired genotype (3, 4). Most applications of this technology have been directed at creating mice with mutations causing loss of function in the gene of interest. The loss of function is most frequently achieved by inserting another gene, such as the neomycin resistance (*neo*^{r}) gene, into an exon of the chosen gene, thereby disrupting the coding sequence of that gene, and at the same time conferring a selectable phenotype (G418^r) on the ES cells that acquire a functional copy of the targeting vector.

A null mutation defines the basal function of a gene and establishes the first critical need for the gene during the embryonic or postnatal periods of the organism. A deeper understanding of the function of a gene is obtained by evaluating the phenotypic consequences generated by a set of mutant alleles. Such mutations could be in the *cis* transcriptional control elements, thereby altering the expression pattern of the gene, or in specific protein regions, thereby revealing functional domains of the gene product. A number of options exist for creating such specific mutations by gene targeting. Two general methods, which make alternative use of insertion or replacement vectors (22) are illustrated in Fig. 1. In both cases, the *neo*^{r} gene is linked to an arbitrary mutation indicated by an asterisk in exon 3. The first scheme requires two steps (6, 13, 17, 24). First, the entire targeting vector is inserted into the cognate chromosomal sequence, generating a duplication of sequences at the endogenous locus. The cells which contain a targeted insertion of the vector are then subjected to a second round of selection to identify subclones in which intrachromosomal homologous recombination has condensed the duplication. During this condensation, the *neo*^{r} gene is removed from the host sequences. Selection for cells in which the reduction has occurred can be accomplished by including in the targeting vector a negative selectable gene, such as the herpes simplex virus thymidine kinase (TK) gene, whose loss renders the cells resistant to gancyclovir (11). The second scheme makes use of a replacement vector to substitute exogenous sequences for endogenous sequences and thereby cotransfers the specific mutation with the *neo*^{r} gene in a single step. There are advantages and disadvantages associated with both schemes. The first scheme requires two separate culture steps. Prolonged in vitro culturing of embryonic stem cells can reduce their ability to contribute efficiently to the germ line. However, this scheme has the advantage that the only modification remaining in the host genome, following reduction of the duplication, is the desired specific mutation. The second scheme involves fewer in vitro manipulations of the pluripotent ES cells, but the genome retains a copy of the *neo*^{r} gene along with the specific mutation. However, since the investigator chooses where to place the *neo*^{r} gene, a potentially neutral site can be selected.

Both schemes rely on the cotransfer of the specific mutation with the positive selectable gene (i.e., *neo*^{r}). The efficiency of this cotransfer is dependent on the position of the crossovers between the targeting vector and chromosomal target locus. In this report, we show that with insertion vectors, the crossovers often occur at some distance from the double-strand break. With replacement vectors, we show that they often occur near the ends of the targeting vector.

MATERIALS AND METHODS

Vectors. The three insertion vectors used in this study were all derived from pIV9.3 (22). pIV9.3-Xh/0.7 contains an 8-bp *Xho*I linker inserted into the *Sca*I site in *Hprt* intron 6; pIV9.3-Xh/2.8 contains an *Xho*I linker inserted into the *Xmn*I site in intron 6; pIV8.1d/4.2 contains a 1.2-kb deletion between two *Sph*I sites in intron 5. The replacement vectors contain 11 kb of *Hprt* sequences extending from an *Eco*RI site in intron 1 to an *Xmn*I site in intron 4, cloned into a pUC-derived plasmid. The *neo*^{r} gene from pMC1Neo (22) was inserted into an *Xho*I site in exon 3 to generate pRV11-1 (*neo*^{r} transcription parallel to *Hprt* transcription) or pRV11-2 (*neo*^{r} transcription in antiparallel orientation). The *Cla*I polymorphisms in both replacement vectors were generated

* Corresponding author.

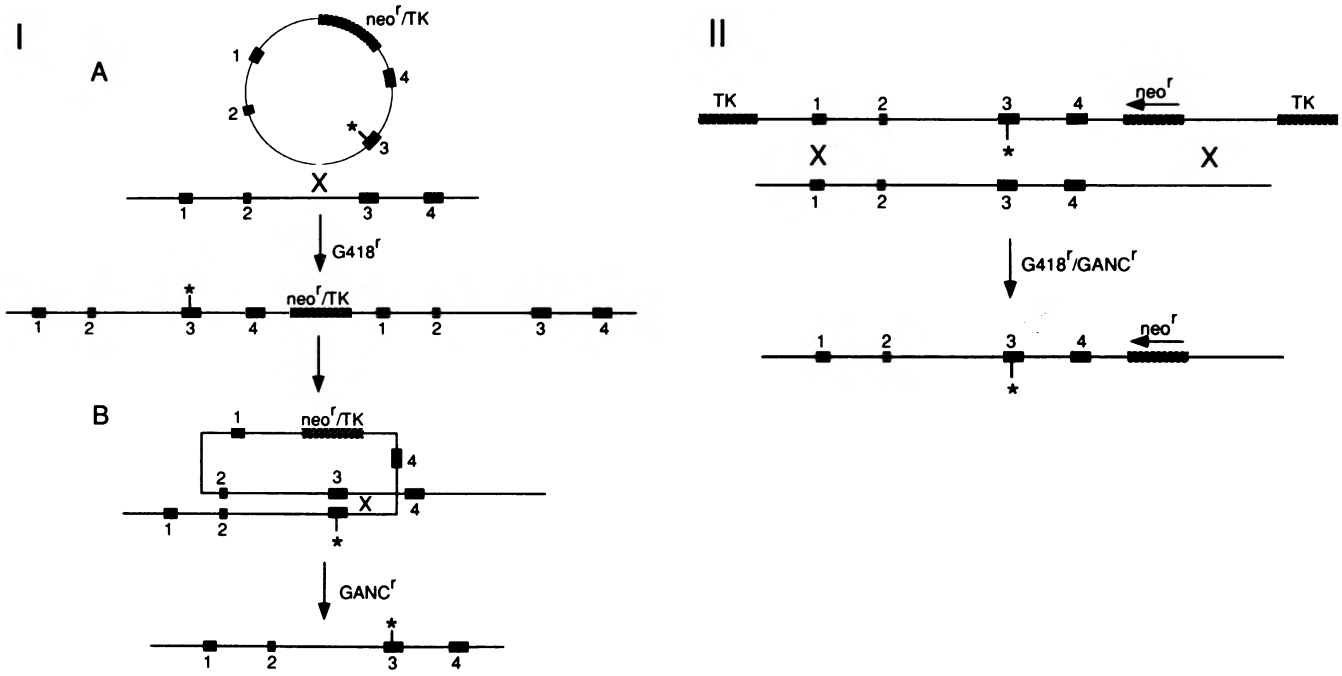

FIG. 1. Targeting a point mutation to a generic gene. (I) Two-step gene replacement using a sequence insertion vector. In the first step (A), an insertion vector (22) containing a point mutation in exon 3 and a linked *neo*^{r} and TK gene cassette pairs and recombines with its target locus, rendering the recipient cell G418^r and GANC^s. In the second step, cells which have lost the TK gene are selected by the presence of gancyclovir. If the loss of the TK-gene is the result of intrachromosomal recombination at the appropriate interval, the resulting GANC^r cell line will have lost the *neo*^{r} TK gene cassette but retained the point mutation. (II) One-step gene replacement using a sequence replacement vector. The replacement vector (top line) contains a point mutation in exon 3, a *neo*^{r} cassette at a neutral position (in this example a 3'-flanking region), and flanking negative selectable markers (HSV TK genes). Recombination with the chromosomal target at positions indicated by the X's will result in transfer of the point mutation and the *neo*^{r} gene and will concomitantly render the cell G418^r and GANC^r. Introns (thin lines), exons (solid boxes), and point mutations (*) are indicated.

by the insertion of *Cla*I linkers at the *Eco*RV sites in introns 1 and 3, thereby destroying the *Eco*RV sites.

Murine *CFTR* sequences were identified in a λ library of genomic DNA from mouse strain 129/Sv/Ev by hybridization to an oligonucleotide containing nucleotides 1537 to 1584 of the human *CFTR* DNA (9). Vector pRVCF11 contains a 12-kb *Bam*HI-to-*Xho*I fragment subcloned from the phage. The *neo*^{r} gene in pRVCF11 contains promoter sequences from the mouse RNA polymerase II gene (nucleotides 1 to 713) (a gift from J. Corden), structural sequences from pMC1Neo (22), and a polyadenylation signal from the murine *Hprt* gene (22). The *neo*^{r} gene was inserted into *CFTR* at a *Kpn*I site in intron 10, 4 kb 3' of exon 10. The phenylalanine residue in exon 10, corresponding to human F-508, was removed by oligonucleotide-directed mutagenesis, using a Bio-Rad reagent kit. A 2.1-kb *Eco*RI fragment containing exon 10 was subcloned into a pBS vector (Stratagene), which was used to generate antisense single-stranded DNA. The mutagenic oligonucleotide 5'-GAAAATATCATTGGT GTTTCC-3' extends from nucleotide 1646 to 1669 from the mouse *CFTR* gene but contains a 3-bp (CTT) deletion between bases 1656 and 1660. Following mutagenesis, clones were screened by DNA sequencing. A fragment containing the appropriate mutation was used to replace the equivalent fragment in the wild-type *CFTR* sequences.

Cell lines and cell culture. The ES cell line, CC1.2 (2), was used in all experiments. Growth and selection conditions were those described previously (11).

Mutant identification. DNA extraction and Southern transfer conditions have been described (22, 23). Identification of the phenylalanine deletion in *CFTR* was performed by a published polymerase chain reaction (PCR) protocol (1) and reagents from Perkin-Elmer Cetus. Denatured cellular DNA (125 ng) was used as the substrate for amplification in a 25-μl reaction mixture containing the following: 2 mM $MgCl_2$, 200 μM each of the four deoxynucleoside triphosphates 0.7 U of *Taq* polymerase; and 300 nM each of the four primers. Amplification consisted of 27 cycles, with each cycle including the denaturation step (95°C, 30 s), annealing step (60°C, 20 s), and extension step (72°C, 60 s). The final cycle included an additional 7-min extension step. Oligonucleotides used to specifically amplify 76 bp of mutant *CFTR* sequences were as follows: 5'-GTTGGCAAGCTTTGACAA CACTCT-3' (antisense nucleotides from exon 10) and 5'-GGTACTATCAAAGAAAATATCATTGG-3' (sense nucleotides from exon 10 missing the critical phenylalanine residue). Also amplified was a 461-bp sequence from *int-2* (12), using oligonucleotides 5'-CTCAGGCGTCCAGTCITG TGG-3' (antisense nucleotides 993 to 973) and 5'-GATACCT CAGGAAGTCCTTCTTC-3' (sense nucleotides 532 to 554). Following amplification, 10 μl of each reaction mixture was analyzed by electrophoresis through a 6% polyacrylamide gel and stained with ethidium bromide.

RESULTS

Point of crossovers with sequence insertion vectors. It is often presumed, by analogy with *Saccharomyces cerevisiae* (8, 14), that when a sequence insertion vector integrates by homologous recombination into the endogenous cognate

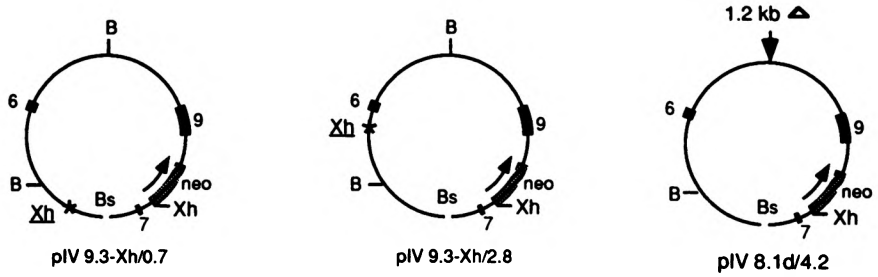

FIG. 2. Targeting vectors. All vectors were derived from pIV9.3 (22), as described in Materials and Methods. Prior to transfer into ES cells, vectors were linearized by digestion with *Bst*EII (Bs). Intron sequences (thin lines), exon sequences (numbered, solid boxes), *neo*^{r} cassette (stippled box), and *Xho*I (Xh) linker insertion sites (*) are indicated. 1.2 kb Δ, 1.2-kb intron deletion; B, *Bgl*II.

locus, the crossovers occur at or near the double-strand break in the vector (7). To evaluate this assumption, a marker was placed at different distances from the site of the double-strand break in the *hprt* insertion vectors shown in Fig. 2. pIV9.3-Xh/0.7 and pIV9.3-Xh/2.8 contain new *Xho*I sites 0.7 and 2.8 kb, respectively, 5′ of the *Bst*EII site used to create the double-strand break, whereas pIV8.1d/4.2 contains a 1.2-kb deletion 4.2 kb 5′ to the *Bst*EII site.

Following the introduction of these targeting vectors into ES cells by electroporation, and selection for colonies capable of growing in medium containing G418 and 6-thioguanine (6-TG), virtually all of the cells were *hprt* mutants as a consequence of targeted disruption of the *Hprt* gene (5, 23). This result is because the ES cells are derived from a male mouse and therefore contain only a single copy of the X-linked *Hprt* gene which confers sensitivity to 6-TG. In addition, the probability of simultaneously acquiring a spontaneous *hprt* mutation (6-TG^r) and a random copy of the *hprt* targeting vector (G418^r) is approximately 4 to 5 orders of magnitude lower than acquiring a targeted *hprt* mutation (22).

To assess the cotransfer of the marker along with the *neo*^{r} gene, independent G418^r 6-TG^r colonies were analyzed by Southern blot hybridization. Figure 3 illustrates a typical Southern blot of DNA from G418^r 6TG^r colonies isolated following the introduction of pIV9.3-Xh/0.7 into ES cells. Three types of events were observed: (i) the faithful cotransfer of the *Xho*I marker with the *neo*^{r} gene, (ii) failure of the *Xho*I marker to be transferred, and (iii) the unexpected transfer of the *Xho*I marker to the 5′ duplication (Fig. 3). The latter event is presumed to occur as a consequence of a gene conversion event, centered around the *Xho*I polymorphism, during the insertion of the targeting vector into the *Hprt* locus. Similar events have been detected in targeting experiments at the CHO *Aprt* locus (15).

Table 1 provides a summary of the fates of the *Xho*I marker following targeted disruption of *Hprt* with insertion vectors pIV9.3-Xh/0.7, pIV9.3-Xh/2.8, and pIV8.1d/4.2. From this table, it is evident that the probability of transferring the marker increases as the distance of the marker from the cut site increases. Thus, a marker 4.2 kb from the double-strand break has a 95% probability of being transferred to the genome, whereas a marker 0.7 kb from the break has only a 26% probability of faithful transfer. These results are consistent with a double-strand break-gap repair model (18, 19, 25) for homologous recombination of insertion vectors and emphasize the fact that the size of the gap at the double-strand break can encompass several kilobases.

Points of crossovers with replacement vectors. Figure 4A illustrates *hprt* replacement vectors that contain the *neo*^{r} gene in exon 3 of *hprt*. Two *Eco*RV restriction endonuclease sites approximately 3 kb on either side of the *neo*^{r} gene were each removed by the addition of 8-bp *Cla*I linkers. The intervals between the polymorphisms and the 5′ and 3′ ends of the linearized vector are designated A and D, respectively, whereas the intervals between the *neo*^{r} gene and the 5′ and 3′ polymorphisms are designated B and C, respectively (Fig. 4). The two vectors pRV11-1 and pRV11-2 differ only in the orientation of the *neo*^{r} gene. Following the introduction of these vectors into ES cells and selection for G418^r 6-TG^r colonies, Southern blot hybridization analysis was used to determine the frequency of transfer of the *Cla*I sites to the endogenous *Hprt* locus (diagrammed in Fig. 4B). Table 2 summarizes these results in terms of whether the crossovers occurred in interval A, B, C, or D. In over 50% of the cell lines examined, recombination occurred in the distal intervals (A and D), thus transferring the *Cla*I polymorphism with the *neo*^{r} gene. The orientation of *neo*^{r} gene transcription in the targeting vector does not appear to influence the intervals in which crossovers occur.

Use of replacement vectors to generate specific mutations. The use of sequence insertion vectors to generate specific mutations in ES cells via scheme I (Fig. 1) has been described (6, 24). Here we illustrate the use of replacement vectors to introduce the most common allele among humans with cystic fibrosis, the Δ*F508* mutation, into the *CFTR* gene in mouse ES cells.

The Δ*F508* mutation is a deletion of 3 nucleotides from exon 10 of the human *CFTR* gene that results in the deletion of a single phenylalanine residue from the mature *CFTR* protein (9, 16). The murine *CFTR* gene shares 78% amino acid identity with its human counterpart, including 84% identity in those amino acids encoded by exon 10 (20, 26). The mouse exon 10 sequence contains a phenylalanine which corresponds to the phenylalanine residue that is deleted in the human Δ*F508* mutation and is flanked by 28 amino acids identical to those in the human exon. Such a striking similarity at the protein sequence level suggests that the deletion of the murine equivalent of Δ*F508* may provide a useful animal model for the most prevalent form of cystic fibrosis.

pRVCF11, the vector that was used to introduce the Δ*F508* mutation into the mouse genome, is shown in Fig. 5A. It contains 12 kb of contiguous mouse genomic DNA, including *CFTR* sequences from intron 9, exon 10, and intron 10. Three coding strand nucleotides, CTT, corresponding to those absent in human Δ*F508*, were deleted from exon 10. A *neo*^{r} gene was inserted in noncoding sequences 3′ to exon 10 to serve as a positive selectable marker for cells transformed by the vector. Several factors were considered

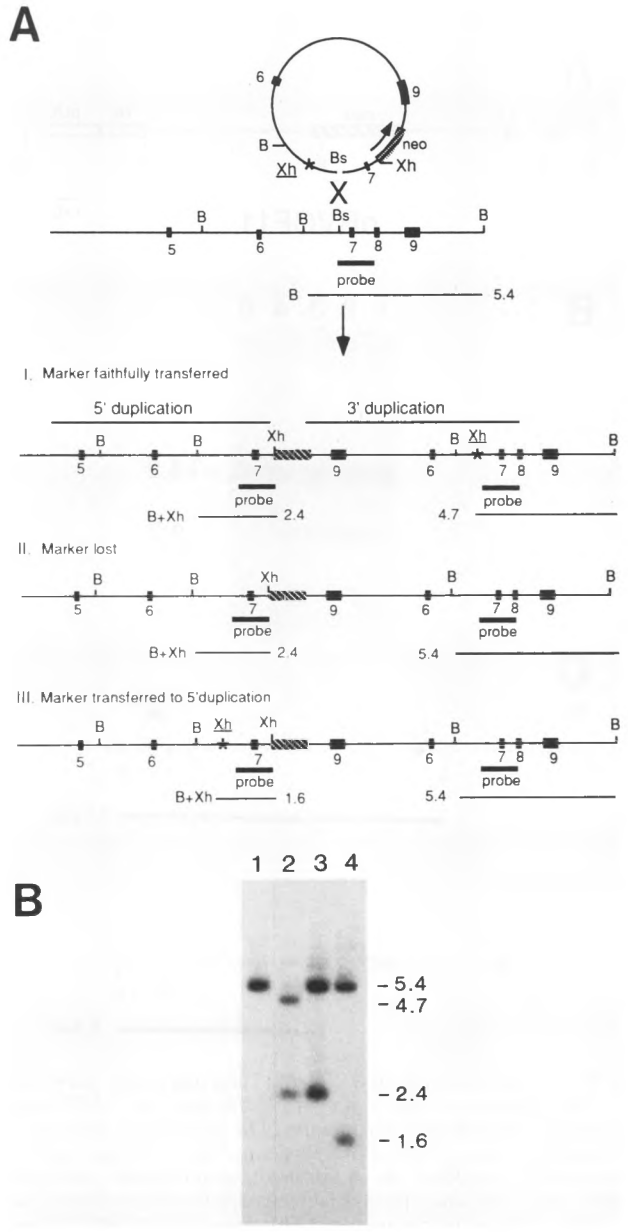

FIG. 3. Transfer of a marker using a sequence insertion vector. (A) Schematic representation of gene targeting. In the top panel, vector pIV9.3-Xh/0.7 is shown paired with the chromosomal *Hprt* locus. The vector is marked with an *Xho*I linker insertion (*Xh*) and the *neo*^{r} cassette. The target locus contains a 5.4-kb *Bgl*II (B) restriction enzyme fragment which hybridizes with the indicated probe. Cells transfected with pIV9.3-Xh/0.7 were selected with G418 and 6-TG to identify those cells in which the vector was inserted at *Hprt*. The fate of the *Xho* linker in the recombinants is followed by digestion with *Bgl*II and *Xho*I (B+Xh) and Southern transfer analysis. Three scenarios are diagrammed: scenario I, the marker is faithfully transferred, resulting in two diagnostic fragments, 4.7 and 2.4 kb in length; scenario II, the marker is lost, resulting in fragments 5.4 and 2.4 kb in length; and scenario III, the marker is transferred to the 5' half of the duplication, resulting in fragments 5.4 and 1.6 kb in length. (B) Southern transfer illustrating the three transfer events. DNA purified from G418^r 6-TG^r cells was digested with *Xho*I and *Bgl*II. The indicated fragment of *Hprt* DNA was used as a probe. Lane 1, DNA from the parental cell line CC1.2; lanes 2, 3, and 4, DNA from transfer events of class I, II, and III, respectively. The sizes (in kilobases) of the hybridizing fragments are indicated to the right.

TABLE 1. Fates of linked markers following targeted disruption with sequence insertion vectors

Vector	No. of clones analyzed	No. of (%) of clones with marker:		
		Faithfully introduced	Lost	Transferred to 5' duplication
pIV9.3-Xh/0.7	23	6 (26)	14 (61)	3 (13)
pIV9.3-Xh/2.8	22	11 (50)	9 (41)	2 (9)
pIV8.1d/4.2	22	21 (95)	1 (5)	0

in choosing the position of the *neo*^{r} gene in order to minimize potential influences that this exogenous gene could have on normal expression of *CFTR*. First, the gene was placed in the transcriptional orientation opposite that of the *CFTR* gene, so that normal *CFTR* transcription would not be terminated prematurely by signals contained in the *neo*^{r}

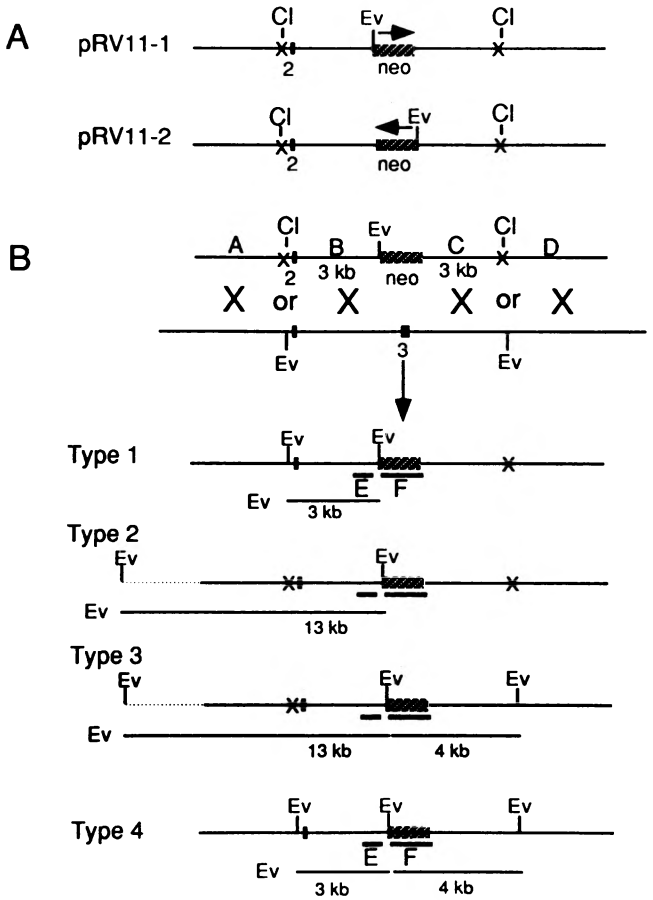

FIG. 4. Transfer of markers using sequence replacement vectors. (A) Vectors pRV11-1 and pRV11-2 contain DNA including *Hprt* exons 2 and 3 with a *neo*^{r} insertion in exon 3. The direction of transcription of the *neo*^{r} gene (depicted by arrows) is opposite in the two vectors. Two *Eco*RV sites (Ev), one on either side of exon 3 in the genomic sequences, have been replaced in the vector by *Cla*I (Cl) linkers. (B) Transfer of the *neo*^{r} gene from the vector (upper line) to the genome (lower line) can occur by two recombination events, one at interval A or B, the other at interval C or D. The four scenarios are indicated along with the diagnostic restriction enzyme (Ev) patterns: type 1, intervals B and D; type 2, intervals A and D; type 3, intervals A and C; and type 4, intervals B and C. E and F are the two probes used for analysis.

TABLE 2. Crossover sites in homologous recombination mediated by replacement vectors

Vector	No. of clones analyzed	No. (%) of clones with the following crossover site:			
		5′ arm		3′ arm	
		A	B	C	D
pRV11-1	41	32 (78)	9 (22)	21 (51)	20 (49)
pRV11-2	31	21 (68)	10 (32)	16 (52)	15 (48)

gene. Second, the *neo*^{r} coding sequences were followed by a strong polyadenylation signal (derived from the murine *Hprt* gene) to reduce the chance of antisense transcription from proceeding into *CFTR* coding sequences. Third, the *neo*^{r} gene was placed 4 kb downstream of the exon 10 splice donor site to minimize any potential interference with excision of the intron 10. Also included in the vector were two copies of the thymidine kinase gene from herpes simplex virus to provide selection markers against cells transformed by random integration of the targeting vector (11).

Following introduction of this vector by electroporation into the ES cell line, CC1.2, clones were selected for resistance to G418 and FIAU. From 2×10^7 cells surviving electroporation, approximately 200 were both G418^r and FIAU^r. DNA purified from 96 of these doubly resistant cell lines was then screened by Southern blot hybridization analysis for the presence of the *neo*^{r} gene inserted in one of the endogenous *CFTR* loci. A total of 13 cell lines carried a *neo*^{r} gene targeted to the correct position adjacent to *CFTR* exon 10. An example of this analysis is shown in Fig. 5B. DNA from the parental cell line, CC1.2, and four G418^r FIAU^r cell lines was digested with *Bam*HI and hybridized to *CFTR* sequences located outside the targeting vector. The probe hybridizes to a 13.2-kb *Bam*HI fragment in wild-type DNA. Cell lines containing the targeted *neo*^{r} gene, however, have an additional *Bam*HI site, and thus the probe reveals a *CFTR* fragment 8.5 kb in length. The presence of the *neo*^{r} gene at the predicted position was also confirmed by four additional enzyme digestions, as well as by using a probe with sequences both internal to and 5′ of the targeting vector. Of the 13 targeted cell lines, all contained the predicted structures, indicating that no additional rearrangements had occurred (data not shown).

To identify those cells in which the Δ*F508* mutation had been cotransferred with the *neo*^{r} gene, a PCR analysis for the presence of the 3-bp deletion was performed. A primer set similar to that used to screen for human Δ*F508* carriers (1) was used to amplify a 76-bp sequence of exon 10. The 5′ primer is elongated only when hybridized with the mutant *CFTR* allele. The positive control in these analyses is the amplification of a 461-bp sequence from the *int-2* gene. The results of such an analysis are shown in Fig. 6. In lane 2 is DNA amplified from the parental cell line CC1.2. Lane 3 contains CC1.2 DNA mixed with cloned *CFTR* DNA containing the Δ*F508* mutation at a molar ratio of 0.6 with respect to the wild-type (CC1.2) *CFTR* sequences. Lanes 4 through 13 contain DNA from 10 G418^r FIAU^r cell lines transformed with pRVCF11. As an internal control, the DNA in lane 7 is from a cell line containing a nontargeted, randomly integrated (and therefore Δ*F508*-containing) copy of the targeting vector. The remaining lanes contain DNA from cells that contain the targeted *neo*^{r} gene. Those DNAs in lanes 4, 5, 8, 11, 12, and 13 all score positive for deletion

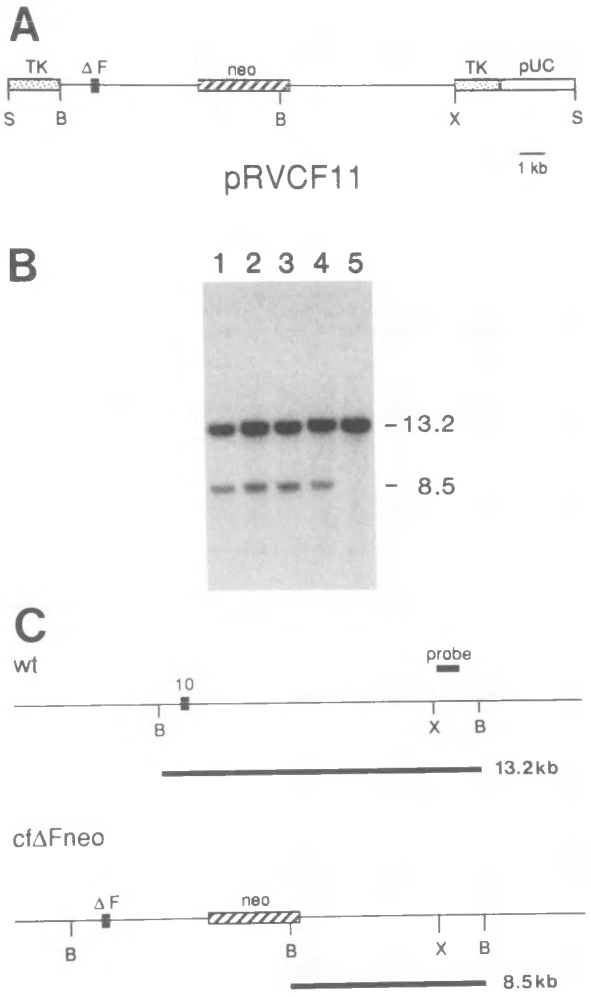

FIG. 5. Gene targeting of *CFTR*. (A) Targeting vector, pRVCF11. A 12-kb fragment of mouse genomic DNA from the *CFTR* locus, extending from a *Bam*HI site in intron 9 to an *Xho*I site in intron 10, was inserted into a modified pUC9 vector. The 3.2-kb *neo*^{r} cassette was inserted at a *Kpn*I site in intron 10, in antiparallel orientation relative to *CFTR*, and a phenylalanine codon was deleted from exon 10. The *CFTR* sequences are flanked by copies of the HSV TK gene. Prior to transfer to ES cells, the vector was linearized at the unique *Sal*I site. (B) Southern transfer analysis of targeted cell lines. DNA was extracted from G418^r FIAU^r cell lines transfected with pRVCF11. Following digestion with *Bam*HI, 5 μg of DNA was electrophoresed through agarose, transferred to nitrocellulose, and probed with a radiolabelled fragment of intron 10 DNA derived from sequences adjacent to those in the vector. Lanes 1 to 4, DNA from four independently isolated G418^r FIAU^r cell lines; lane 5, DNA from the parental ES cell line, CC1.2. The band containing the 13.2-kb wild-type allele is more intense than the band containing the 8.5-kb mutant allele, because it also contains DNA from the wild-type *CFTR* locus contributed by the STO fibroblast feeder cells upon which the ES cells are grown. (C) Schematic representation of the Southern transfer data. The wt line shows the DNA from the wild-type *CFTR* locus containing the 13.2-kb *Bam*HI fragment hybridizing to the intron 10 probe. The cfΔFneo line indicates the same locus containing both targeted modifications, the Δ*F* deletion and the *neo*^{r} insertion; B, *Bam*HI; S, *Sal*I, X, *Xho*I. Intron sequences (thin lines), exon 10 (solid box), *neo*^{r} gene (hatched box), HSV TK gene (stippled box); pUC9 (open box) are indicated.

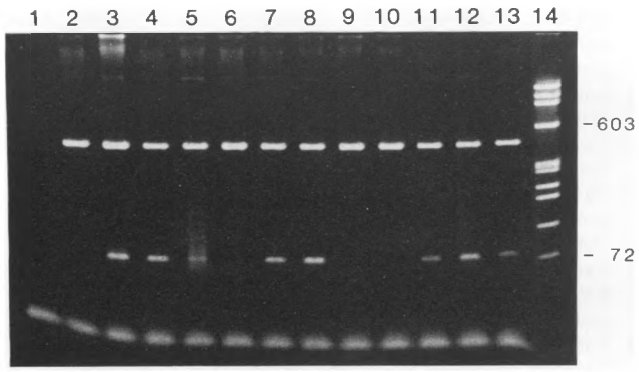

FIG. 6. Identification of the ΔF mutation. DNA extracted from individual ES cell lines was amplified by PCR as described in Materials and Methods. The reaction products were electrophoresed through 6% polyacrylamide and stained with ethidium bromide. Lane 1, no cell DNA; lane 2, DNA from the parental ES cell line, CC1.2; lane 3, CC1.2 DNA mixed with pRVCF11 DNA at a molar ratio of 0.6 copy of pRVCF11 per genomic copy of *CFTR*; lanes 4 to 13, DNA from G418^r GANC^r cells transfected with pRVCF11 (DNA in lane 7 contains a randomly integrated copy of pRVCF11, and DNAs in the remaining lanes contain targeted copies of the *neo*^{r} gene); lane 14, *Hae*III digest of ϕX174 DNA. The positions of the 603- and 72-bp ϕX174 fragments are indicated.

of the phenylalanine residue. Further confirmation of the presence of the deletion was obtained by cloning and sequencing the mutant allele from two of the $\Delta F508$ cell lines (data not shown). A summary of these results is as follows: from a total of 96 G418^r FIAU^r cells, 13 were shown to contain the *neo*^{r} gene targeted to one of the *CFTR* loci. Of the 13 targeted cell lines, 6 exhibited cotransfer of the $\Delta F508$ mutation, a frequency of nearly 50%. Four of these cell lines have been used to generate chimeric animals, and the DNAs from two lines have been successfully passed through the germ line. It should be noted that ES cell lines that contain the *neo*^{r} gene in exon 10, but not the $\Delta F508$ mutation, can be used to generate control mice to determine whether the presence of the *neo*^{r} gene itself has any detectable phenotypic consequences. The targeting frequency at *CFTR* reported here is significantly higher than that reported previously (10) using a similar, replacement-type vector. One explanation for this difference is that the vector DNA used in this study is isogenic with that of the target, a situation shown to maximize targeting frequencies (5, 21).

DISCUSSION

The analysis of the cotransfer of restriction enzyme polymorphisms with a selectable genetic marker has allowed us to determine the regions of crossover between targeting vectors and their chromosomal targets. The frequency of transfer of the nonselected marker from the vector to the target was shown to be dependent upon both the position of the marker relative to the ends of the linear targeting vector and upon the topology of the vector.

When insertion vectors were used as substrates, the probability of marker transfer could be correlated with the distance of that marker from the double-strand break in the vector. A marker located 0.7 kb from the cleavage site was faithfully transferred only 26% of the time, while a marker 4.2 kb from the break was transferred at a frequency of 95%. The relatively low frequency of cotransfer of the end-proximal marker suggests that gene conversion events, emanating from the double-strand break, may be occurring. Such events would be predicted if recombination occurred via the double-strand break repair model for vector integration (18, 19).

The crossovers associated with replacement vector-mediated gene targeting were shown to occur towards the ends of the targeting vector. Thus, nonselectable markers as far as 3 kb from the positive selection marker could be cotransferred at frequencies greater than 50%. The utility of such a cotransfer was demonstrated by the cotransfer of the $\Delta F508$ mutation of the *CFTR* locus with a *neo* gene located 4 kb from the mutation in a neighboring intron sequence.

In addition to their bearing on potential recombination mechanisms, these observations have implications for the construction of targeting vectors designed to introduce subtle mutations to specific chromosomal sites. When insertion vectors are used to perform two-step gene replacement, the unselected marker should be placed at some distance from the site of the double-strand break. On the other hand, when replacement vectors are used to cotransfer the nonselectable marker, the position of the selectable marker relative to the nonselectable mutation is of less importance.

ACKNOWLEDGMENTS

We thank Marjorie Allen, Sheila Barnett, and Carol Lenz for excellent technical assistance with cell culture and Linda Oswald for help with the preparation of the manuscript.

REFERENCES

1. **Ballabio, A., R. A. Gibbs, and C. T. Caskey.** 1990. PCR test for cystic fibrosis deletion. Nature (London) **343**:220. (Letter.)
2. **Bradley, A., M. Evans, M. H. Kaufman, and E. Robertson.** 1984. Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature (London) **309**:255–256.
3. **Capecchi, M. R.** 1989. Altering the genome by homologous recombination. Science **244**:1288–1292.
4. **Capecchi, M. R.** 1989. The new mouse genetics: altering the genome by gene targeting. Trends Genet. **5**:70–76.
5. **Deng, C., and M. R. Capecchi.** 1992. Reexamination of gene targeting frequency as a function of the extent of homology between the targeting vector and the target locus. Mol. Cell. Biol. **12**:3365–3371.
6. **Hasty, P., R. Ramírez-Solis, R. Krumlauf, and A. Bradley.** 1991. Introduction of a subtle mutation into the *Hox-2.6* locus in embryonic stem cells. Nature (London) **350**:243–246.
7. **Hasty, P., J. Rivera-Pérez, C. Chang, and A. Bradley.** 1991. Target frequency and integration pattern for insertion and replacement vectors in embryonic stem cells. Mol. Cell. Biol. **11**:4509–4517.
8. **Hinnen, A., J. B. Hicks, and G. R. Fink.** 1978. Transformation of yeast. Proc. Natl. Acad. Sci. USA **75**:1929–1933.
9. **Kerem, B., J. M. Rommens, J. A. Buchanan, D. Markiewicz, T. K. Cox, A. Chakravarti, M. Buchwald, and L.-C. Tsui.** 1989. Identification of the cystic fibrosis gene: genetic analysis. Science **245**:1073–1080.
10. **Koller, B. H., H.-S. Kim, A. M. Latour, K. Brigman, R. C. Boucher, Jr., P. Scambler, B. Wainwright, and O. Smithies.** 1991. Toward an animal model of cystic fibrosis: targeted interruption of exon 10 of the cystic fibrosis transmembrane regulator gene in embryonic stem cells. Proc. Natl. Acad. Sci. USA **88**:10730–10734.
11. **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.
12. **Moore, R., G. Casey, S. Brookes, M. Dixon, G. Peters, and C. Dickson.** 1986. Sequence, topography and protein coding potential of mouse *int-2*: a putative oncogene activated by mouse mammary tumour virus. EMBO J. **5**:919–924.
13. **Morse, M. L., E. M. Lederberg, and J. Lederberg.** 1956.

Transductional heterogenotes in *E. coli*. Genetics **41**:758–779.
14. **Orr-Weaver, T. L., J. W. Szostak, and R. J. Rothstein.** 1981. Yeast transformation: a model system for the study of recombination. Proc. Natl. Acad. Sci. USA **78**:6354–6358.
15. **Pennington, S. L., and J. H. Wilson.** 1991. Gene targeting in Chinese hamster ovary cells is conservative. Proc. Natl. Acad. Sci. USA **88**:9498–9502.
16. **Riordan, J. R., J. M. Rommens, B. Kerem, N. Alon, R. Rozmahel, Z. Grzelczak, J. Zielenski, S. Lok, N. Plavsic, J.-L. Chou, M. L. Drumm, M. C. Iannuzzi, F. S. Collins, and L.-C. Tsui.** 1989. Identification of the cystic fibrosis gene: cloning and characterization of complementary DNA. Science **245**:1066–1073.
17. **Scherer, S., and R. W. Davis.** 1979. Replacement of chromosome segments with altered DNA sequences constructed *in vitro*. Proc. Natl. Acad. Sci. USA **76**:4951–4955.
18. **Sun, H., D. Treco, and J. W. Szostak.** 1991. Extensive 3′-overhanging, single-stranded DNA associated with the meiosis-specific double-strand breaks at the *ARG4* recombination initiation site. Cell **64**:1155–1161.
19. **Szostak, J. W., T. L. Orr-Weaver, R. J. Rothstein, and F. W. Stahl.** 1983. The double-strand-break repair model for recombination. Cell **33**:25–35.
20. **Tata, F., P. Stanier, C. Wicking, S. Halford, H. Kruyer, N. J. Lench, P. J. Scambler, C. Hansen, J. C. Braman, R. Williamson, and B. J. Wainwright.** 1991. Cloning the mouse homolog of the human cystic fibrosis transmembrane conductance regulator gene. Genomics **10**:301–307.
21. **teRiele, H., E. R. Maandag, and A. Berns.** 1992. Highly efficient gene targeting in embryonic stem cells through homologous recombination with isogenic DNA constructs. Proc. Natl. Acad. Sci. USA **89**:5128–5132.
22. **Thomas, K. R., and M. R. Capecchi.** 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell **51**:503–512.
23. **Thomas, K. R., C. Deng, and M. R. Capecchi.** 1992. High-fidelity gene targeting in embryonic stem cells by using sequence replacement vectors. Mol. Cell. Biol. **12**:2919–2923.
24. **Valancius, V., and O. Smithies.** 1991. Testing an "in-out" targeting procedure for making subtle genomic modifications in mouse embryonic stem cells. Mol. Cell. Biol. **11**:1402–1408.
25. **Valancius, V., and O. Smithies.** 1991. Double-strand gap repair in a mammalian gene targeting reaction. Mol. Cell. Biol. **11**: 4389–4397.
26. **Yorifuji, T., W. K. Lemna, C. F. Ballard, C. L. Rosenbloom, R. Rozmahel, N. Plavsic, L.-C. Tsui, and A. L. Beaudet.** 1991. Molecular cloning and sequence analysis of the murine cDNA for the cystic fibrosis transmembrane conductance regulator. Genomics **10**:547–550.