Targeting of Nonexpressed Genes in Embryonic Stem Cells Via Homologous Recombination

RANDALL S. JOHNSON, MORGAN SHENG, MICHAEL E. GREENBERG, RICHARD D. KOLODNER, VIRGINIA E. PAPAIOANNOU, BRUCE M. SPIEGELMAN*

Gene targeting via homologous recombination-mediated disruption in murine embryonic stem (ES) cells has been described for a number of different genes expressed in these cells; it has not been reported for any nonexpressed genes. Pluripotent stem cell lines were isolated with homologously recombined insertions at three different loci: c-*fos*, which is expressed at a low level in ES cells, and two genes, adipsin and adipocyte P2 (aP2), which are transcribed specifically in adipose cells and are not expressed at detectable levels in ES cells. The frequencies at which homologous recombination events occurred did not correlate with levels of expression of the targeted genes, but did occur at rates comparable to those previously reported for genes that are actively expressed in ES cells. Injection of successfully targeted cells into mouse blastocysts resulted in the formation of chimeric mice. These studies demonstrate the feasibility of altering genes in ES cells that are expressed in a tissue-specific manner in the mouse, in order to study their function at later developmental stages.

GENE TARGETING IN ES CELLS HAS been envisioned as a method for studying the effects of specific mutations in the mouse (*1–4*). Genes of interest would be altered by homologous recombination-mediated disruption after DNA transfection of ES cells. These ES cell lines containing altered genes would then be used to create chimeric mice, through injection of the cells into blastocysts. If the germ line is colonized by ES cells in chimeric mice, subsequent breeding would allow the generation of mouse strains that are heterozygous or homozygous for the altered gene. This protocol appears feasible because of two advances: the development of techniques that allow the growth and maintenance of blastocyst-derived cells (ES cells) in a pluripotent state (*5–7*) and the ability to select for homologous recombination-mediated gene replacement in mammalian cell culture (*1, 8–10*).

There have been suggestions that the frequency of a homologous recombination event may depend on the transcription rate of the gene concerned (*3, 9*), although experiments with a human bladder carcinoma cell line indicated that homologous recombination at a nonexpressed locus was possible (*10*). If rates of homologous recombination in ES cells are strongly correlated with rates of expression, it might restrict gene targeting to those genes that are expressed ubiquitously or are active in early development. We report here that two genes expressed in adult mouse adipose tissue, but not detectably expressed in ES cells, can be readily disrupted by homologous recombination.

The adipsin and aP2 genes are both transcriptionally activated during adipocyte differentiation (*11–14*) and the mRNAs encoding the proteins are expressed predominantly in adipose tissue in the mouse (*15–17*). To determine the relative expression of these mRNAs in ES cells, we performed Northern blotting with the ES parent cell line CC1.2 (*5*) and several subclones. No mRNA encoded by either of these genes could be detected, even on prolonged exposures of the autoradiographs (Fig. 1, B and C). In contrast, c-*fos* is expressed in these cells, albeit at a low level compared to the nerve growth factor (NGF)–induced PC-12 cells (*18*) used as a control (Fig. 1A).

Gene targeting vectors (Fig. 2) were constructed to permit the selection of gene replacement events by a procedure similar to that outlined by Mansour *et al.* (*9*). Vectors contained the thymidine kinase (TK) gene of herpes simplex virus (HSV)–1 (*19*). The entire vector sequence is expected to be retained after random integration into the genome, making the cell sensitive to nucleoside analogs, such as gancyclovir (*20*), through the action of the HSV-TK gene product. Homologous recombination with an endogenous sequence is expected to result in the loss of nonhomologous sequences, allowing the cells to retain their native resistance to gancyclovir (*9, 20*). Thus one can use gancyclovir to select against the presence of the TK gene product and in so doing select against random integrants and enrich for homologous recombinants.

Neomycin-resistance gene (*neo*^{r}) cassettes (*1*) were inserted into genomic clones so as to disrupt transcription or translation, or both, and render the genes dysfunctional. The neomycin-resistance marker also allows a positive genetic selection for cells that have stably incorporated the vector into the genome, through resistance to the drug G418. The vectors were linearized and transfected into ES cells by electroporation. After selection with G418 and gancyclovir, colonies resistant to both drugs were clonally expanded, and DNA was isolated for

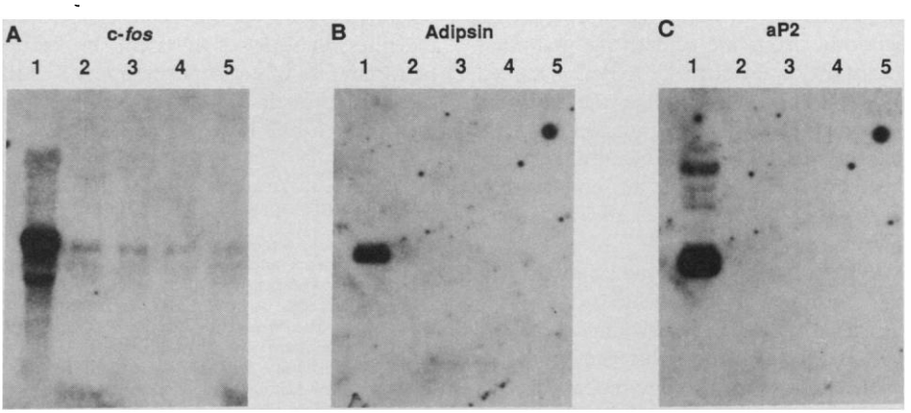

Fig. 1. Expression of c-*fos*, adipsin, and aP2 mRNA in ES cells. Blots were probed with (A) c-*fos*–, (B) adipsin (*25*)–, and (C) aP2 (*25*)–labeled cDNAs. RNA was isolated as described (*25*) and 10 μg of total cytoplasmic RNA was loaded per lane. Electrophoresis, blotting, and probing was performed as previously described (*25*). Adipsin and aP2 blots were overexposed to show the lack of a signal from ES cell lines. (**A**) Lane 1, PC-12 cells stimulated by NGF for 30 min (positive control for c-*fos* expression) (*18*); lane 2, CC1.2, parent ES cell line; lane 3, c-*fos* homologous recombinant f-16; lane 4, adipsin homologous recombinant a-16; and lane 5, aP2 homologous recombinant 8-3. (**B**) Lane 1, differentiated 3T3-F442A adipocytes; lane 2, CC1.2; lane 3, f-16; lane 4, a-16; and lane 5, 8-3. (**C**) Lanes were the same as in (B).

R. S. Johnson, R. D. Kolodner, B. M. Spiegelman, Dana-Farber Cancer Institute and the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115.
M. Sheng and M. E. Greenberg, Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, MA 02115.
V. E. Papaioannou, Department of Pathology, Tufts University Schools of Medicine and of Veterinary Medicine, Boston, MA 02111.

*To whom correspondence should be addressed.

Southern blotting (Fig. 3).

G418-resistant cells occurred at a relatively constant frequency among the different DNA constructs transfected (approximately 0.1% of the transfected cells) (Table 1). The number of homologous recombinants, as determined by Southern blotting of genomic DNA, was between 5 to 10% of the colonies resistant to both G418 and gancyclovir, although the total number of doubly resistant colonies varied somewhat among constructs. Rates of homologous recombination, expressed as the number of homologous recombinants over the number of stably transfected (G418-resistant) cells, were 5×10^{-5} for c-*fos*, 4×10^{-5} for adipsin, and 3×10^{-4} for aP2. Although these rates are based on a relatively small number of recombinants (15 in all), they are all comparable to those for an expressed gene [that is, 10^{-3} to 2.5×10^{-5} for hypoxanthine-guanine phosphoribosyl-transferase (HPRT)] (*1*).

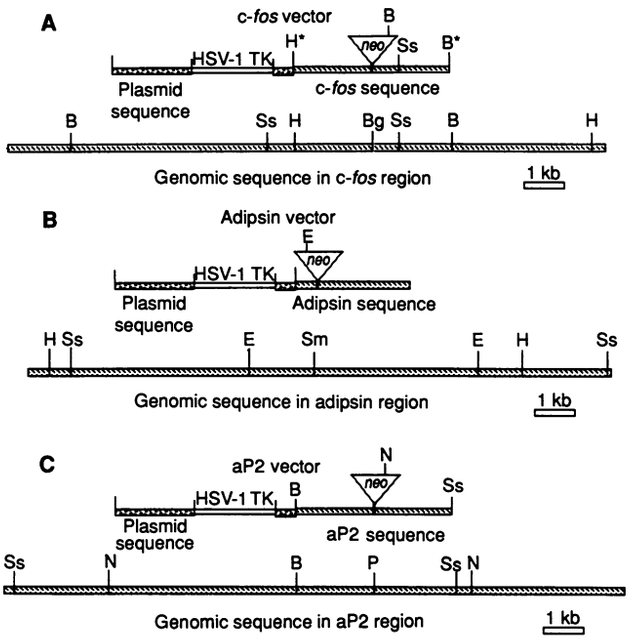

Fig. 2 Transfection vectors containing the disrupted c-*fos*, adipsin, and aP2 gene sequences. (**A**) The parent plasmid vector used to construct all of the above vectors contained the Bam HI 3.4-kb fragment of HSV-1 and its TK gene cloned into the Nae I site of the pBluescript SK(+) plasmid. The genomic sequences described above were cloned into the polylinker of this plasmid. The *neo*^{r} used for gene disruption was derived from the Xho I–Sal I fragment of pMCIneo/polyA+ (*1*). B, Bam HI; Bg, Bgl I; E, Eco RI; H, Hind III; N, Nco I; P, Pst I; Sm, Sma I; and Ss, Sst I. The c-*fos* vector was derived from Hind III–Bam HI genomic fragment containing c-*fos*. The asterisk indicates the genomic sites that were lost during cloning into the plasmid vector. The c-*fos* gene was interrupted by insertion of the *neo*^{r} cassette at a Bgl I site in the first exon. Before transfection the vector was linearized at a unique Not I site in the plasmid polylinker, with linearization giving the orientation shown at the top of the figure. (**B**) The adipsin vector was derived by Bal 31 digestion of an Eco RI genomic fragment. The adipsin gene was interrupted in the third exon at a unique Sma I site and linearized and transfected as above. (**C**) The aP2 vector was derived from a 4.5-kb Bam HI–Sst I genomic fragment. The aP2 gene was interrupted after the transcription start site and before a translation start site at a Pst I site.

The ES cell lines with genes targeted by homologous recombination were analyzed for expression of adipsin, aP2, and c-*fos* mRNAs to ensure that the selection process did not enrich for cells with aberrant patterns of expression for those genes. The targeted cell lines and the parent cell line have very similar patterns of expression for the three genes, with no evidence of adipsin or aP2 mRNA (Fig. 1, lanes 2 to 5). Although the possibility cannot be ruled out that aP2 and adipsin are transcribed at very low levels (*21*), they appear to be good examples of genes for which no expression is anticipated or detected in ES cells. There is no reason to suspect that targeting these loci is any more or less difficult than targeting other loci not detectably expressed in these cells. This belief is supported by two observations: (i) the rates of homologous recombination for aP2 and adipsin are comparable to reported rates of recombination for other, expressed genes (*1, 9, 22, 23*), and (ii) we see no obvious relation between expression and targeting rates for a transcribed gene (c-*fos*) and for two genes not detectably expressed in our experiments.

Chimeric mice have been derived in outbred (CD1) or inbred (C57BL/6J, AG/CamPa) mouse strains, or both, with separate ES cell lines containing disrupted aP2, adipsin, and c-*fos* genes; the parent cell line and the clones used to make these mice were karyotyped to ascertain stable, diploid chromosome numbers. The mice are now being bred to ascertain germ line transmission of the altered genes (*5*). The ability of the cells to contribute to tissue formation in chimeric mice indicates that the cells have retained pluripotency during the selection process (*5–7*).

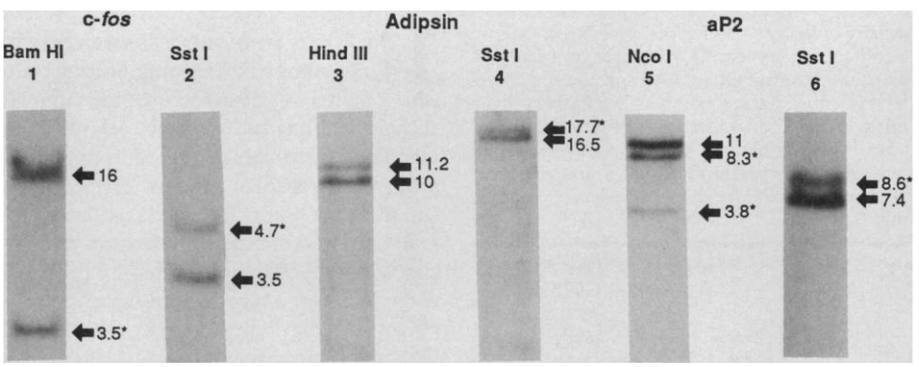

Fig. 3. Identification of ES clones with homologously recombined insertions by DNA blotting. ES cell lines resistant to both G418 and gancyclovir were selected as described in Table 1. DNA from cell lines was isolated, digested with restriction enzymes, blotted, and probed as described (*25*). The asterisk indicates bands generated by homologous recombination events. Numbers refer to kilobase pairs of DNA. Lanes 1 and 2, DNA from the c-*fos* homologous recombinant f-16 probed with the c-*fos* genomic sequence between the Bgl I site in the first exon and the Sph I site in the third intron. Lane 1, Bam HI digestion. Lane 2, Sst I digestion. Lane 1 shows the 16-kb fragment resulting from the hybridization of the probe with the unaltered copy of the gene and the 3.5-kb fragment resulting from the cleavage within the gene as a result of a site in the *neo*^{r}. Lane 2 shows the unaltered fragment at 3.5-kb as well as the 1.2-kb larger fragment at 4.7 kb engendered by the *neo*^{r} cassette. Lanes 3 and 4, DNA from the adipsin homologous recombinant clone a-16 probed with adipsin cDNA (*25*). Lane 3, digested with Hind III. Lane 4, digested with Sst I. In both lanes the lower band corresponds to the remaining copy of the endogenous gene; the upper band is the same fragment plus the 1.2-kb *neo*^{r} insertion. In lane 3 the unaltered fragment is at 10 kb and the gene plus the *neo*^{r} insertion is 11.2 kb. In lane 4 the unaltered gene is at 16.5 kb and the gene plus insertion is at 17.7 kb. Lanes 5 and 6, DNA from the aP2 homologous recombinant clone 8-3 probed with aP2 cDNA (*25*). Lane 5, digested with Nco I. Lane 6, digested with Sst I. The Nco I fragment is split because of an Nco I site in *neo*^{r}, giving the 3.8-kb and 8.3-kb fragments. The 11-kb fragment corresponds to the remaining unaltered copy of the gene. Lane 6 shows the endogenous Sst I fragment at 7.4 kb plus a 1.2-kb larger fragment at 8.6 kb containing the *neo*^{r} insertion. Blots were also probed with the *neo*^{r} sequence to confirm that the homologous recombinant bands contained *neo*^{r} cassette insertions (data not shown).

The three genes we have targeted in these experiments are likely to be important in mouse development and physiology, but their functions remain undefined genetically. The adipsin gene encodes a serine protease with complement factor D activity and has links to systemic energy balance and obesity (*15, 24*); the aP2 gene product is an adipocyte-specific fatty acid binding protein (*25–27*) whose precise role in physiology is unknown. The c-*fos* proto-oncogene is involved in the regulation of gene transcription (*28, 29*), and disruption of this gene may have numerous effects on cell function. The analysis of the phenotypic effect of disruption of these three genes should lead to a greater understanding of their function in the organism.

It is probable that many factors are involved in determining the frequency at which detectable homologous recombination occurs at a given genetic locus (*2, 3*). Our data, however, demonstrate that it is feasible to target genes not expressed in ES cells in order to study their function in mouse development and physiology.

Table 1. Rates of homologous recombination after double drug selection. Electroporations were performed with the linearized vector described in Fig. 2. The data are pooled from several transfections. ES cells were grown in Dulbecco's modified Eagle's medium (DMEM) with 15% fetal bovine serum (FBS) and 0.1 m*M* 2-mercaptoethanol on gamma-irradiated (4000 rads) Sto feeder layers made G418-resistant by transfection with the vector pSV2neo. Cells (2×10^7) in 1 ml of medium were electroporated at 750 V/cm in DMEM and FBS with vector DNA at 1 n*M*. These cells were plated out on G418-resistant feeder layers. The number of viable cells 24 hours after electroporation was always between 40 to 60%. At 36 hours after transfection, G418 (100 μg/ml) and gancyclovir (2 μ*M*) were added. After selection for 10 days, drug-resistant colonies were picked and expanded on Sto feeder layers without further drug selection. The asterisk indicates that one control plate for each transfection was selected with G418 alone and was plated at 1/20 the density of doubly selected plates (5×10^5 cells on a 150-cm^2 plate versus 2×10^7 cells on a 150-cm^2 plate), and the number of G418-resistant colonies for the entire transfection was extrapolated from this number. Clones listed as homologous recombinants were confirmed as such by digestion with two or more restriction enzymes and Southern blotting, as shown for the representative clones in Fig. 3.

DNA transfected	Cells electroporated	No. G418* resistant	No. G418 and gancyclovir resistant	Homologous recombinants
c-*fos* vector	4×10^7	2×10^4	16	1
Adipsin vector	10^8	5×10^4	22	2
aP2 vector	8×10^7	4×10^4	142	12

REFERENCES AND NOTES

1. K. R. Thomas and M. R. Capecchi, *Cell* **51**, 503 (1987).
2. M. R. Capecchi, *Science* **244**, 1288 (1989).
3. M. Frohman and G. Martin, *Cell* **56**, 145 (1989).
4. V. E. Papaioannou *et al.*, *J. Embryol. Exp. Morphol.* **44**, 93 (1978).
5. A. Bradley *et al.*, *Nature* **309**, 255 (1984).
6. M. J. Evans and M. H. Kaufman, *ibid.* **292**, 154 (1981).
7. G. R. Martin, *Proc. Natl. Acad. Sci. U.S.A* **78**, 7634 (1981).
8. K. R. Thomas *et al.*, *Cell* **44**, 419 (1986).
9. S. L. Mansour *et al.*, *Nature* **336**, 348 (1988).
10. O. Smithies *et al.*, *ibid.* **317**, 230 (1985).
11. K. S. Cook, C. R. Hunt, B. M. Spiegelman, *J. Cell. Biol.* **100**, 514 (1985).
12. D. A. Bernlohr, M. A. Bolanowski, T. J. Kelly, Jr., M. D. Lane, *J. Biol. Chem.* **260**, 5563 (1985).
13. P. Djian, M. Phillips, H. Green, *J. Cell. Physiol.* **124**, 554 (1985).
14. A. Doglio, C. Davi, P. Grimaldi, G. Ailhaud, *Biochem. J.* **238**, 123 (1986).
15. K. S. Cook, D. L. Groves, H. Y. Min, B. M. Spiegelman, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 6480 (1985).
16. K. M. Zezulak and H. Green, *Mol. Cell. Biol.* **5**, 419 (1985).
17. D. A. Bernlohr *et al.*, *Biochem. Biophys. Res. Commun.* **132**, 850 (1985).
18. M. Sheng *et al.*, *Mol. Cell. Biol.* **8**, 2787 (1988).
19. S. McKnight, *Nucleic Acids Res.* **8**, 5949 (1980).
20. M. H. St. Clair, C. U. Lambe, P. A. Furman, *Antimicrob. Agents Chemother.* **31**, 844 (1987).
21. G. Sarkar and S. S. Sommer, *Science* **244**, 331 (1989).
22. T. Doetschman, N. Maeda, O. Smithies, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8583 (1988).
23. A. Joyner *et al.*, *Nature* **338**, 153 (1989).
24. B. S. Rosen *et al.*, *Science* **244**, 1483 (1989).
25. B. M. Spiegelman, M. Frank, H. Green, *J. Biol. Chem.* **258**, 10083 (1983).
26. V. Matarese and D. Bernlohr, *ibid.* **263**, 14544 (1988).
27. R. C. Hresko *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 8835 (1988).
28. R. J. Distel *et al.*, *Cell* **49**, 835 (1987).
29. T. Curran and B. R. Franza, *ibid.* **55**, 395 (1988).

30. We thank A. Bradley for the CC1.2 ES cell line; J. Spencer for technical assistance; and B. Satterberg, M. Mercola, M. Musacchio, L. Michalowsky, L. Reid, B. Koller, O. Smithies, J. Rossant, R. Kucherlapati, D. Coen, and E. Robertson for helpful discussions. Supported by a grant from NIH (DK 31405). R.J. and M.S. were supported by Lucille P. Markey predoctoral fellowships. B.S. is an Established Investigator of the American Heart Association.

3 July 1989; accepted 2 August 1989

Foregut Fermentation in the Hoatzin, a Neotropical Leaf-Eating Bird

ALEJANDRO GRAJAL, STUART D. STRAHL, RODRIGO PARRA, MARIA GLORIA DOMINGUEZ, ALFREDO NEHER

The only known case of an avian digestive system with active foregut fermentation is reported for the hoatzin (*Opisthocomus hoazin*), one of the world's few obligate folivorous (leaf-eating) birds. Hoatzins are one of the smallest endotherms with this form of digestion. Foregut fermentation in a flying bird may be explained by increased digestive efficiency by selection of highly fermentable and extremely patchy resources, coupled with microbial nutritional products and secondary compound detoxification. This unexpected digestive system gives a new perspective to the understanding of size limitations of vertebrate herbivores and to the evolution of foregut fermentation.

FOREGUT MICROBIAL FERMENTATION as a means of digesting fibrous plant matter has been reported in mammals such as ruminants, monkeys, sloths, and macropodid marsupials (*1*). Although a few bird species display hindgut fermentation (*2*), there are no documented cases of extensive foregut fermentation structures or associated digestive physiology in the entire class Aves. We now report a well-developed ruminant-like digestive system in a neotropical folivorous bird, the hoatzin, *Opisthocomus hoazin*. This is the first report of this digestive system outside the mammals, and opens new insights into the evolution of foregut fermentation.

The hoatzin is a 750-g cuculiform bird that ranges from the Guianas to Brazil and inhabits riverine swamps, gallery forests, and oxbow lakes (*3, 4*). Early descriptions suggested that the crop of this species has replaced the gizzard and proventriculus as the primary site of digestion (*5*). None of these authors documented or suggested foregut fermentation, although some noted that the characteristic odor of the bird was similar to fresh cow manure (*6*).

At our study site (*7*), more than 80% of the hoatzin's diet is composed of green leaves. Although the birds fed on the leaves of 52 species of plants in 25 families, 90% of the diet is composed of only 17 plant spe-

A. Grajal, Department of Zoology, University of Florida, Gainesville, FL 32611.
S. D. Strahl, Wildlife Conservation International, New York Zoological Society, Bronx, NY 10460.
R. Parra and A. Neher, Instituto de Producción Animal, Universidad Central de Venezuela, Facultad de Agronomía, Maracay, Estado Aragua, Venezuela.
M. G. Dominguez, Rowett Research Institute, Bucksburn, Aberdeen, Scotland, United Kingdom.