868–872 Nucleic Acids Research, 1997, Vol. 25, No. 4 © 1997 Oxford University Press

Targeted integration of DNA using mutant *lox* sites in embryonic stem cells

Kimi Araki[1], Masatake Araki[2] and Ken-ichi Yamamura[1,2,*]

[1]Department of Developmental Genetics, Institute of Molecular Embryology and Genetics, Kumamoto University School of Medicine, Kumamoto 862, Japan and [2]Gene Technology Center, Kumamoto University, Kumamoto 862, Japan

Received September 27, 1996; Revised and Accepted December 13, 1996

ABSTRACT

Site-directed DNA integration has been achieved by using a pair of mutant *lox* sites, a right element (RE) mutant *lox* site and a left element (LE) mutant *lox* site [Albert *et al.* (1995) *Plant J.*, 7, 649–659], in mouse embryonic stem (ES) cells. We established ES cell lines carrying a single copy of the wild-type *lox*P or LE mutant *lox* site as a target and examined the frequency of site-specific integration of a targeting vector carrying a *lox*P or RE mutant *lox* site induced by Cre transient expression. Since our targeting vector contains a complete *neo* gene, random integrants can form colonies as in the case of a gene targeting event through homologous recombination. With our system, the frequency of site-specific integration via the mutant *lox* sites reached a maximum of 16%. In contrast, the wild-type *lox*P sites yielded very low frequencies (<0.5%) of site-specific integration events. This mutated *lox* system will be useful for 'knock-in' integration of DNA in ES cells.

INTRODUCTION

The Cre–*lox* recombination system of bacteriophage P1 is currently the most powerful tool for genetic manipulation both *in vitro* (1–5) and *in vivo* (6–9). Cre recombinase catalyzes reciprocal site-specific recombination between two *lox*P sites. Consequently, Cre mediates both intramolecular (excisive or inversional) and intermolecular (integrative) recombination. In integrative recombination, a circular DNA carrying a *lox*P site is inserted into a *lox*P site on a chromosome. However, this integration reaction is quite inefficient, because the integrated DNA, which has *lox*P sites at both ends, is easily removed again though excisive recombination if the Cre recombinase is still present (Fig. 1A). Therefore, a special selection system in which only targeted integrants can survive is indispensable for targeted integration into *lox*P sites. For example, Sauer and Henderson (1) used a promoter fusion strategy. In this system, the targeting vector contains a *lox*P site followed by a promoterless *tk* gene and the chromosomal target is a *lox*P site located 3′ of a promoter. Only upon *lox*P-specific targeted integration does the promoterless *tk* gene on the targeting vector fuse to the chromosomally placed promoter, thereby making the cells HAT resistant. Fukushige and Sauer (2) further improved this strategy using an ATG-less *lox–neo* translational fusion system. In this system, the targeting vector contains a promoter followed by ATG–*lox* and the chromosomal target is a *lox*P site located 5′ of the ATG-less *neo* gene. Only upon site-specific recombination dose the promoter–ATG–*lox* element on the targeting vector fuse with the ATG-less *neo* gene, resulting in the generation of a functional *lox–neo* fusion gene to confer G418 resistance on the cells.

Recently, Albert *et al.* (10) devised a new strategy different from the methods mentioned above. They identified three sets of mutant *lox* sites that favor integrative recombination over the excisive reaction. The *lox*P site is composed of an asymmetric 8 bp spacer flanked by 13 bp inverted repeats. They introduced nucleotide changes into the left 13 bp element (LE mutant *lox* site) or the right 13 bp element (RE mutant *lox* site) (Fig. 1C). Recombination between the LE mutant *lox* site and the RE mutant *lox* site produces the wild-type *lox*P site and a LE+RE mutant site that is poorly recognized by Cre, resulting in stable integration (Fig. 1B and C).

This LE/RE mutant system has many potential uses, because it does not need any special system for selecting targeted integration. Only a mutant *lox* site is needed as the chromosomal target. However, the efficiency of targeted integrative recombination in the LE/RE mutant system in mammalian cells is not known. Albert *et al.* used plant, not mammalian, cells and since they utilized the promoter fusion strategy as well as the LE/RE mutant system, they could not precisely determine the efficiency of recombination in the LE/RE mutant system. In this study we have used embryonic stem (ES) cells, which are extensively used in genome engineering and adopted a simple selection system to examine the frequency of targeted insertion over random integration. We show here that the LE/RE mutant system is more effective than wild-type *lox*P and that the efficiency of targeted integration was in the range 2–16%.

MATERIALS AND METHODS

Plasmids

The *lox*P, *lox*66 and *lox*71 sequences (10) were synthesized and cloned into pBluescript (pBS)-SK$^-$ (Stratagene). The pCAGloxPbsr and pCAGlox71bsr plasmids were constructed by first

*To whom correspondence should be addressed. Tel: +81 96 373 5316; Fax: +81 96 373 5321; Email: yamamura@gpo.kumamoto-u.ac.jp

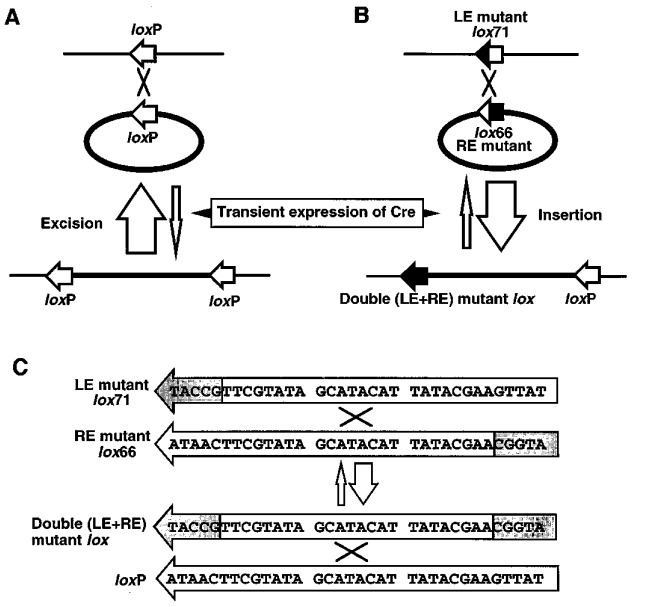

Figure 1. The LE/RE mutant system. Horizontal arrows represent *lox* sites. Vertical arrows indicate the direction and frequency of the reaction. (**A**) Recombination between *lox*P sites. The integrated DNA flanked by two *lox*P sites is easily removed again if Cre recombinase is present, therefore, the insertion product is unstable during transient Cre expression. (**B**) Recombination between a LE mutant site and a RE mutant site produces a wild-type *lox*P site and a double (LE+RE) mutant site. Since the LE+RE mutant site exhibits reduced binding affinity for Cre recombinase, the inserted DNA is stable. The filled parts of the *lox* arrows indicate mutated sequences. (**C**) Nucleotide sequences of *lox* sites. The shaded sequences indicate nucleotide base changes of the mutated *lox* sites.

introducing a *bsr Hin*dIII fragment from pSV2bsr (Kaken-Seiyaku, Tokyo) (11) into *lox*P- or *lox*71-containing plasmids. Then, the *lox*P–*bsr* or *lox*71–*bsr* fragment was ligated into pBS-CAGG, which was constructed by inserting a *Sal*I–*Pst*I fragment from pCAGGS (12) into pBS-SK$^-$.

Plasmid ploxPNZneo and plox66NZneo were assembled from components of pBSnlslacZ, pMC1NeopolyA (Stratagene) and *lox* sequences. Plasmid pBSnlslacZ contains the *lacZ* gene fused with the nuclear localization signal (NLS) derived from the SV40 large T gene and an intron and poly(A) signal derived from SV40 (13). The MC1neopolyA fragment was inserted into the 3′-end of NLS–*lacZ* and the *lox*P or *lox*66 fragment was inserted into the 5′-end of NLS–*lacZ*.

The Cre expression vector, pCAGGS-Cre, was constructed by inserting the *cre* fragment from pBS185 (Life Technologies) into pCAGGS (12).

Cell culture and electroporation

The ES cell line TT2 (14) was grown as described (15) except for the use of G418-resistant primary mouse embryo fibroblasts as feeder layers.

In the case of electroporation with the pCAGloxPbsr or pCAGlox71bsr plasmids, 10 µg *Spe*I-digested DNA and 1×10^7 cells were used. The cells were electroporated using a BioRad Gene Pulser and after 48 h they were fed with medium supplemented with 4 µg/ml Blasticidin S (Kaken-Seiyaku, Tokyo). Selection was maintained for 5 days and then colonies were picked into 24-well plates and expanded for freezing. The clones were analyzed by Southern blotting to select cell lines showing a single copy integration pattern.

For the electroporation experiments designed to detect targeted integration into the *lox* site, ploxPNZneo, plox66NZneo and pCAGGS-Cre were used in their circular forms. The electroporated cells were selected with G418 at 200 µg/ml for 1 week. Colonies were stained with X-gal as described (16) or picked and expanded for DNA analysis.

Southern hybridization, PCR and DNA sequencing

Six micrograms of genomic DNA were digested with appropriate restriction enzymes, electrophoresed on a 0.9% agarose gel and then blotted onto a nylon membrane (Boehringer Mannheim). Hybridization was performed using a DIG DNA Labeling and Detection Kit (Boehringer Mannheim). For PCR analysis, DNAs (0.1–0.5 µg) were subjected to 28 cycles of amplification (each cycle consisting of 1 min at 94°C, 2 min at 55°C and 2 min at 72°C) with a thermal cycler. The 5′-primers and 3′-primers were AG2 (5′-CTGCTAACCATGTTCATGCC-3′) and LZUS3 (5′-GCGCATCGTAACCGTGCAT-3′) for amplification of the 5′-junction fragment and T7 (5′-AATACGACTCAGTATAG-3′) and BSR1 (5′-CTTCTCTGTCGCTACTTCTAC-3′) for amplification of the 3′-junction fragment. One half of the PCR reaction product was loaded and analyzed on an agarose gel.

The nucleotide sequences of the regions surrounding the *lox* sites derived from PCR from the targeted clones were determined with a thermo sequenase fluorescent labeled primer cycle sequencing kit (US Biochemicals).

RESULTS

Experimental design

The experimental design outlined in Figure 2 was used to assess the efficiency of targeted recombination between the mutant lox sequences, *lox*71 and *lox*66 (Fig. 1), which we chose from the three sets of mutants reported by Albert *et al.* (10). We established six ES cell lines carrying a single copy of *lox*P (loxP-5 and loxP-6) or *lox*71 (lox71-2, lox71-19, lox71-20 and lox71-22). The *lox* site was placed between the CAG promoter and the *bsr* gene as a target. Since the transformants were selected with Blasticidin S (i.e. *bsr* gene expression), the CAG promoter was active in the cell lines. After introducing the targeting plasmids, ploxPNZneo and plox66NZneo, we could detect *lox* site-mediated integration by monitoring expression of the inserted gene, i.e. the *lacZ* gene fused with the NLS derived from the SV40 large T gene (NLS–*lacZ*) in this study. The targeting plasmids also contained the complete *neo* gene with a promoter and poly(A) signal (see Fig. 2). In this system, although both random and targeted integrants become G418 resistant, only the targeted integrants are stained blue by X-gal and the percentage of blue colonies indicates the frequency of targeted integration. It is expected that in the case of site-specific recombination between wild-type *lox*P sites, the inserted vector will be excised by the remaining Cre recombinase activity. On the other hand, targeted integration via *lox*66 and *lox*71 will be stable because of the poor affinity of the LE+RE mutant site for the Cre recombinase.

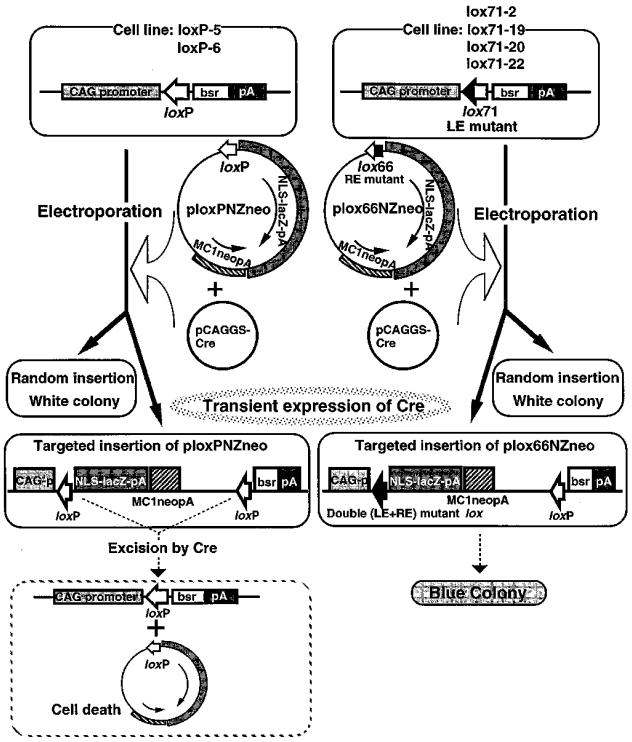

Figure 2. Description of the experiment. Cell lines carrying a single copy of the CAG promoter–*lox*–*bsr*–*pA* fragment were established. loxP-5 and loxP-6 carry a wild-type *lox*P and lox71-2, lox71-19, lox71-20 and lox71-22 carry a LE mutant *lox*71. The large arrows represent *lox* sequences and the black parts of the arrows represent mutations. The cell lines were co-electroporated with the targeting vector, ploxPNZneo or plox66NZneo, and the Cre expression vector, pCAGGS-Cre, and then selected with G418. In random integration, colonies are not stained with X-gal, because there is no promoter sequence for the NLS–*lacZ* gene. In targeted integration, colonies are stained for NLS–*lacZ* expression, because the NLS–*lacZ* gene is joined to the CAG promoter. In the case of recombination between *lox*P sites, the inserted vector sequence flanked by wild-type *lox*P sites is easily eliminated by the remaining Cre activity, resulting in cell death through G418 selection. On the other hand, in the case of recombination between *lox*71 and *lox*66, a double mutant *lox* site (represented by a black arrow), which is poorly recognized by Cre, is generated, resulting in stable integration that confers NLS–*lacZ* expression on the cells.

Targeted recombination frequency

The cell lines carrying a *lox*P or *lox*71 site were co-electroporated with a constant amount (20 µg) of ploxPNZneo or plox66NZneo and various amounts of the Cre expression vector, pCAGGS-Cre, as indicated in Table 1. After G418 selection for 1 week, the colonies were stained with X-gal and then scored. All of the positive colonies turned blue within 2 h, reflecting the strong activity of the CAG promoter. The blue colonies derived from the same parental cell line showed almost the same intensity of staining, suggesting that no 'gene-trap'-type integration had occurred. However, the staining intensities varied among different parental lines, probably due to the different promoter activity in each parental cell line. The results are summarized in Table 1. The targeting frequencies for normal *lox*P sites were very low (<0.5%) irrespective of the amount of pCAGGS-Cre. We could find two blue colonies at most. On the other hand, the frequencies of targeted integration between *lox*66 and *lox*71 were in the range 2–16% with 20 µg each of pCAGGS-Cre and plox66NZneo. These results indicate that the LE/RE mutant system is more efficient as to Cre-mediated site-specific integration. When we introduced plox66NZneo into the *lox*P-carrying cell lines or ploxPNZneo into the *lox*71-carrying cell lines, the frequencies were low (0.3–1%), although they were slightly better than in the case of *lox*P–*lox*P recombination. This demonstrates that the combination of the RE and LE mutants is important for achieving targeted integration.

Table 1. Cre-mediated DNA targeting into wild or mutant *lox* sites

Cell line	Transfected DNA: Amount of pCAGGS-Cre (µg) + ploxPNZneo (20µg)	Transfected DNA: Amount of pCAGGS-Cre (µg) + plox66NZneo (20µg)	Plated cell number (x10^6)	Number of G418^r colonies	Number of Blue colonies	Targeting frequency (% of blue/ total)
loxP-5	0		0.6	139	0	0
	10		3.0	270	0	0
	20		7.2	1254	3	0.2
	40		1.2	218	1	>0.5
		20	6	962	3	0.3
loxP-6	0		3.8	492	0	0
	2		3.8	432	1	>0.2
	10		2	186	0	0
	20		5.2	834	1	>0.2
	40		1.6	257	1	>0.4
		20	4	684	3	0.4
lox71-2		0	8	1242	0	0
		10	9	1552	21	1.4
		20	8	818	17	2.1
		40	8	908	13	1.4
lox71-19		0	0.6	54	0	0
		10	3	270	16	6.0
		20	5.4	695	114	16
		30	6	485	64	13
		40	1.6	174	0	0
	20		4	375	4	1.1
lox71-20		0	2.3	177	0	0
		2	2.3	224	6	2.7
		10	3.5	218	11	5.0
		20	4.8	636	52	8.2
		30	8	402	18	4.5
		40	1.6	225	16	7.1
	20		3	341	2	0.6
lox71-22		0	7	1056	0	0
		10	8	1289	34	2.6
		20	8	1008	32	3.2
		40	8	897	24	2.7

We found differences in the targeting frequency among the four cell lines carrying *lox*71. Cell line lox71-19 showed 8-fold higher efficiency than cell line lox71-2. These results indicate that the targeting efficiency may depend on the chromosomal position of the integration site. Interestingly, the positive colonies derived from cell line lox71-19 showed the strongest X-gal staining among the four lines (data not shown), suggesting the possibility that the *lox*71 target site is located at an 'active' chromosomal position. Concerning the positive colonies derived from the *lox*P-carrying cell line, their staining intensities were similar to those derived from cell lines lox71-20 and lox71-22. Therefore, the low targeting efficiency of *lox*P sites is not due to a position effect.

Analysis of integration into lox71 targets

To confirm site-directed integration of the vector into chromosomal *lox* targets, we picked 48 colonies from among the transformants of cell line lox71-19 electroporated with 20 µg each of pCAGGS-Cre and plox66NZneo. Five out of the 48 clones showed positive X-gal staining. We prepared genomic DNAs from these five clones and also from seven negative clones and analyzed them by Southern blotting using pBS as a probe. As shown in Figure 3A, targeted integration should give a 4.3 kb

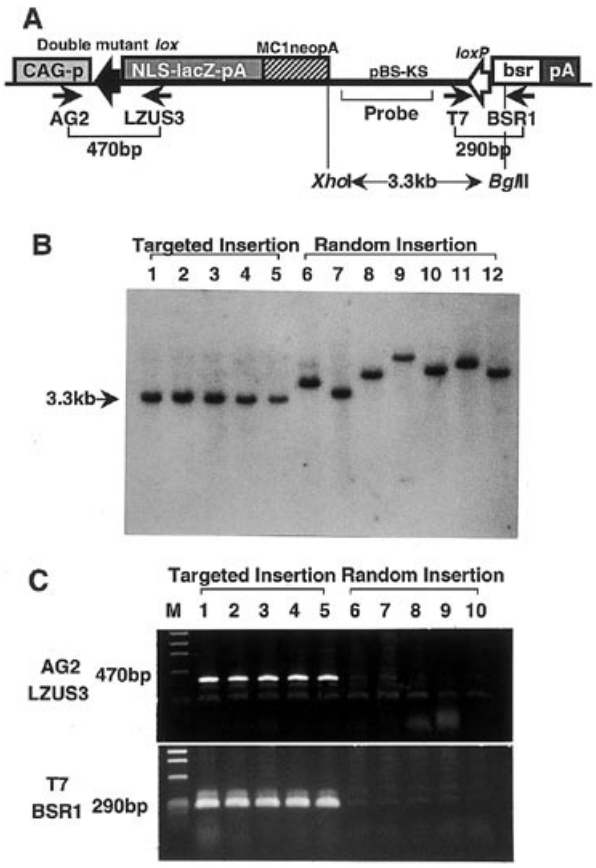

Figure 3. PCR and Southern analyses of genomic DNA from transformants of lox71-19. (**A**) The expected DNA structure for site-specific integration. The small arrows indicate the positions and directions of the PCR primers. The expected sizes of PCR products are indicated. (**B**) Southern blot analysis of *Xho*I + *Bgl*II-digested genomic DNAs from: lanes 1–5, X-gal-positive clones; lanes 6–12, X-gal-negative clones. The probe used for this Southern analysis is indicated in (A). (**C**) PCR analysis. The samples were the same as lanes 1–10 in (B). The primers used are indicated to the left of the photograph.

band, whereas bands of various sizes are expected in the case of random integration. In agreement with this expectation, all the five X-gal-positive clones gave a 4.3 kb band (Fig. 3B, lanes 1–5), in contrast to the 'white' clones (lanes 6–12). Since there were no extra bands other than the 4.3 kb one, a single copy of the vector was inserted in each case. PCR analyses were also performed for confirmation. The junction of the integration can be amplified using the primer pairs AG2 and LZUS3 for the 5′-junction and T7 and BSR1 for the 3′-junction (see Fig. 3A). As shown in Figure 3C, only the X-gal-positive clones gave a band of the expected size. Furthermore, we determined the nucleotide sequence surrounding the *lox* sites of the PCR products and confirmed generation of the wild-type *lox*P site and the double mutant *lox* site (data not shown). These results demonstrate that the clones stained with X-gal are targeted integrants and that the recombination occurred between *lox* sites.

DISCUSSION

We show here that the mutant *lox* sites promoted site-specific integration events much more frequently than those obtained with wild-type *lox* sites in ES cells. The reason for this clear-cut result is probably the use of a simple selection system. It allows us to score the efficiency of targeted integration over random integration by counting blue colonies. We believe this is the first report showing successful Cre-mediated site-specific targeting without selection of targeting events. The highest frequency of site-specific integration was 16% with cell line lox71-19. This frequency is comparable to that of 'knock-out' gene targeting, in which negative selection or promoter trapping is not involved. Therefore, this LE/RE mutant system is a practical method for genetic manipulation in ES cells.

The efficiency of Cre-mediated, targeted integration depends on the amount of Cre expression vector and thus the level of Cre recombinase activity. In our experiment, the optimal amount was 20 μg and a higher or lower amount of the Cre expression plasmid reduced the targeting efficiency. Baubonis and Sauer (3) reported a similar observation for purified Cre protein in the *lox–neo* fusion system. Since the CAG promoter has strong activity, 3- to 5-fold higher than that of the phosphoglycerate kinase-1 (PGK) promoter (unpublished data), we originally expected that <20 μg would be optimal for targeted recombination. Our results suggest that it is necessary to use a strong promoter for effective insertion. On the other hand, an excess amount of the Cre expression plasmid causes excisive recombination, even when a double mutant *lox* site is used. If a more effective pair of mutant *lox* sequences is found, the targeting frequency might increase.

We observed an 8.5-fold difference in targeting efficiency among the cell lines carrying *lox*71. Since all the cell lines used carried a single copy of the target *lox* site, the chromosomal position effect is considered to be the cause of the difference. Baubonis and Sauer (3) also observed a 50-fold difference in targeting efficiency per target *lox*P site in human osteosarcoma cell lines using purified Cre protein. As they used cell lines carrying different numbers of *lox*P sites, the difference in targeting efficiency could be larger. Our finding that the highest *lacZ*-expressing cell line gave the highest targeting efficiency suggests that targeting efficiency might be proportional to the level of gene expression, although the number of examined cell lines was limited. Increasing the number of parental cell lines could make the cause(s) of the position effects on recombination and gene expression more apparent.

When we introduced plox66NZneo into the *lox*P-carrying cell lines or ploxPNZneo into the *lox*71-carrying cell lines, the frequencies were low, as in the case of *lox*P–*lox*P recombination. This indicates that excisive recombination between the *lox*71 and *lox*P sites occurs at a normal rate.

The advantage of this mutant *lox* system is that only a mutant *lox* site is needed as the chromosomal target. We believe that the method described here will be useful for genetic manipulation in ES cells, including conditional gene targeting and gene trapping, as this system allows site-specific integration of any DNA sequence into a defined *lox* site.

ACKNOWLEDGEMENTS

We wish to thank Y.Kiyonaga for technical assistance and Dr K.Abe for critical reading of the manuscript. This work was supported by grants from the Ministry of Education, Science and Culture, by a grant from the Yamanouchi Foundation for Research on Metabolic Disorders, by a grant from the Osaka Foundation for Promotion of Clinical Immunology and a grant from the Science and Technology Agency.

REFERENCES

1 Sauer,B. and Henderson,N. (1990) *New Biologist*, **2**, 441–449.
2 Fukushige,S. and Sauer,B. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 7905–7909.
3 Baubonis,W. and Sauer,B. (1993) *Nucleic Acids Res.*, **21**, 2025–2029.
4 Smith,A.J.H., De Sousa,M.A., Kwabi-Addo,B., Heppell-Parton,A., Impey,H. and Rabbitts,P. (1995) *Nature Genet.*, **9**, 376–385.
5 Zhang,Y., Riesterer,C., Ayrall,A.-M., Sablitzky,F., Littlewood,T.D. and Reth,M. (1996) *Nucleic Acids Res.*, **24**, 543–548.
6 Lakso,M., Sauer,B., Mosinger,B.,Jr, Lee,E.J., Manning,R.W., Yu,S.H., Mulder,K.L. and Westphal,H. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6232–6236.
7 Orban,P.C., Chui,C. and Marth,J.D. (1992) *Proc. Natl. Acad. Sci. USA*, **89**, 6861–6865.
8 Araki,K., Araki,M., Miyazaki,J. and Vassalli,P. (1995) *Proc. Natl. Acad. Sci. USA*, **92**, 160–164.
9 Kühn,R., Schwenk,F., Aguet,M. and Rajewsky,K. (1995) *Science*, **269**, 1427–1429.
10 Albert,H., Dala,E.C., Lee,E. and Ow,D.W. (1995) *Plant J.*, **7**, 649–659
11 Izumi,M., Miyazawa,H., Kamakura,T., Yamaguchi,I., Endo,T. and Hanaoka,F. (1991) *Exp. Cell Res.*, **197**, 229–233.
12 Niwa,H., Yamamura,K. and Miyazaki,J. (1991) *Gene*, **108**, 193–200.
13 Kimura,S., Niwa,H., Moriyama,M., Araki,K., Abe,K., Miike,T. and Yamamura,K. (1994) *Dev. Growth Differentiat.*, **36**, 521–527.
14 Yagi,T., Tokunaga,T., Furuta,Y., Nada,S., Yoshida,M., Tsukada,T., Saga,Y., Takeda,N., Ikawa,Y. and Aizawa,S. (1993) *Anal. Biochem.*, **214**, 70–76.
15 Niwa,H., Araki,K., Kimura,S., Taniguchi,S., Wakasugi,S. and Yamamura,K. (1993) *J. Biochem.*, **113**, 343–349.
16 Joyner,A.L. (1993) *Gene Targeting: A Practical Approach*. Oxford University Press, Oxford, UK.