Nucleic Acids Research, 1995, Vol. 23, No. 19 4003–4004

A *lacZ*–hygromycin fusion gene and its use in a gene trap vector for marking embryonic stem cells

Dipa Natarajan and Catherine A. Boulter*

Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, UK

Received July 27, 1995; Accepted August 22, 1995

The use of gene trap vectors in embryonic stem (ES) cells and transgenic mice has proved to be a powerful method for identifying new genes involved in mammalian development (1–4). Generally, they have a *lacZ* marker gene downstream of a splice acceptor site and rely on integration into introns for their expression. Not only can insertion of these vectors cause disruption of the gene into which they integrate, but it is also possible to visualize the transcription pattern of the disrupted gene itself, by staining with X-gal for expression of β-galactosidase from the *lacZ* gene. More sophisticated gene trap vectors have now been developed carrying a *lacZ*–neomycin fusion gene, β*geo*, which produces a chimaeric fusion protein with both β-galactosidase and neomycin phosphotransferase activities, conferring the added advantage of selection with G418 for insertion within expressed genes (3). Such vectors can also be used for cell marking experiments by integrating within constitutively active genes. Recently β*geo* gene traps have been successfully used to mark ES cells for *in situ* localization of their derivatives in chimaeric mice, providing a useful tool for mosaic analysis *in vivo* (5).

The usefulness of β*geo* gene trap vectors for cell marking is however limited to cell lines that do not already carry a neomycin phosphotransferase (*neo*) gene and are thus not resistant to G418. As *neo* is widely used as a selectable marker for mammalian cells, this excludes the majority of cells carrying expression constructs, such as ES cell lines that have undergone gene targeting events. We have therefore created a novel gene fusion between the *lacZ* and hygromycin B phosphotransferase (*hyg*) (6) genes, named β*gyg*, and have used it to construct a gene trap vector which carries this gene downstream of the splice acceptor site. Here we show that the fusion protein produced by this vector has both β-galactosidase and hygromycin phosphotransferase activities. Moreover, we have successfully introduced this vector into a G418 resistant ES cell line, SR2-3 (7), and by selecting for hygromycin expression, have isolated marked clones that constitutively express β-galactosidase in all cells both before and after differentiation *in vitro*.

The gene trap vector we have constructed, pβgyg, is based on pSAβgeo, which carries a β*geo* fusion gene downstream of the splice acceptor site from adenovirus major late transcript (3). Essentially, the *neo* gene and poly-A site from pSAβgeo has been replaced with a *hyg* gene and phosphoglycerate kinase (PGK) poly-A site from pKJ23 (M. McBurney, University of Ottawa), such that the *hyg* gene is in frame with the *lacZ* gene to generate the fusion gene β*gyg* (Fig. 1). This was achieved as described in the legend to Figure 1. Sequencing across the junction of the fusion site between the *lacZ* and *hyg* genes confirmed that they were indeed in frame. To determine whether β*gyg* was functional in mammalian cells, 30 μg of pβgyg was linearized by digesting with *Kpn*I, and electroporated into 10^7 SR2-3 ES cells. After 14 days selection with 0.3 mg/ml hygromycin B (Sigma), 22 resistant colonies were obtained, demonstrating that the hygromycin fusion protein retains biological activity. Whilst hygromycin has been reported to be active as an N-terminal fusion protein with thymidine kinase (8), this is, to our knowledge, the first demonstration of its activity as a C-terminal fusion with β-galactosidase. Of the 22 hygromycin resistant clones, 10 (45%) stained strongly with X-gal; of these, seven gave staining in >80% of cells and two in 100% of cells. In both of these clones, named S3 and S21, every cell was strongly positive after staining for 4 h, as shown in Figure 2A for the S3 clone.

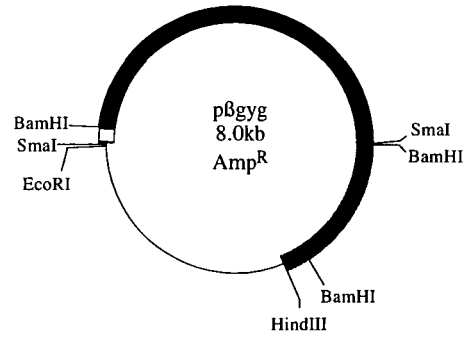

Figure 1. A map of the *lacZ*–hygromycin gene trap vector, pβgyg, showing the *hyg* gene and PGK poly-A site (grey box) fused in frame downstream of the *lacZ* gene (black box). The splice acceptor site from adenovirus major late transcript (white box) lies upstream of the *lacZ* gene. The vector was constructed by digesting pKJ23 (McBurney, University of Ottawa) with *Bam*HI and *Hin*dIII to generate a 1.5 kb fragment containing the *hyg* gene and PGK poly-A site. This was ligated between the *Bam*HI and *Hin*dIII sites of pUC18 to generate pUChyg. A 3.8 kb *Sma*I fragment from pSAβgeo (3) containing the splice acceptor site and *lacZ* gene was then inserted into the *Sma*I site in pUChyg at the 5′ end of the *hyg* gene. The nucleotide sequence of the fusion junction between the *lacZ* and *hyg* genes is as follows: 5′-TGT cag ggg atc ccc cgg gga tcc gcc acc atg gct AAA-3′. The codons in upper case letters are the codons encoding amino acid 1021 in β-galactosidase and amino acid 2 in hygromycin phosphotransferase.

* To whom correspondence should be addressed

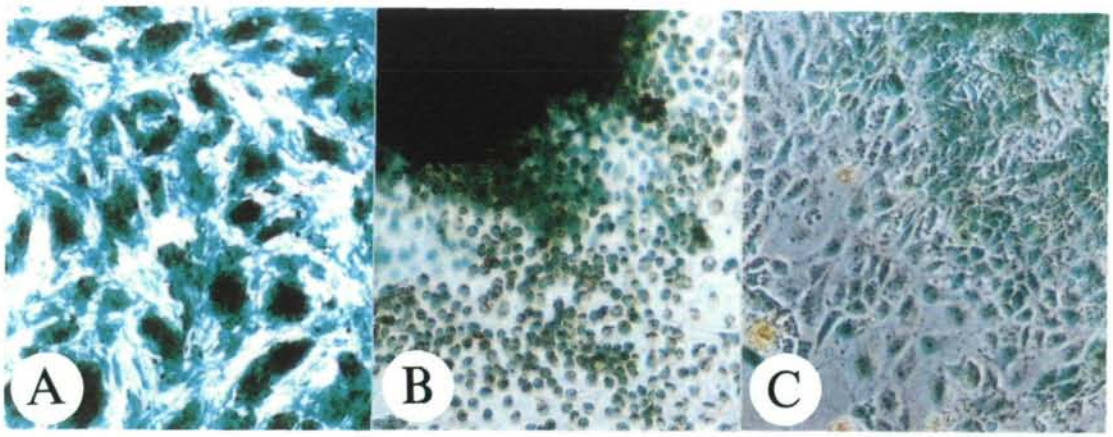

Figure 2. Expression of β-galactosidase in the S3 ES cell clone. (**A**) All the stem cells in the culture are strongly positive for expression, after staining with X-gal for 4 h. (**B** and **C**) Maintenance of β-galactosidase expression in all cells after differentiation *in vitro*, 2 days (B) and 6 days (C) after plating the embryoid bodies. Magnification is ×125.

To determine whether expression of the β*gyg* gene was maintained on cell differentiation, embryoid bodies were derived from the S3 clone, as described (9). After culturing for 10 days in bacteriological dishes, cystic embryoid bodies were obtained. Several days after these were replated onto gelatinized tissue culture plates, a wide variety of differentiated cell types were observed in the cultures, including endoderm, neurones and cardiac muscle. On staining with X-gal, all cells were positive for β-galactosidase expression (Fig. 2B and C). This result demonstrates that the pβgyg gene trap vector can be used to mark ES cells and that constitutive expression of the β-galactosidase–hygromycin fusion protein is maintained on their differentiation.

ACKNOWLEDGEMENTS

We would like to thank Drs M. McBurney and G. Friedrich for kind gifts of the plasmids pKJ23 and pSAβgeo, respectively, and acknowledge the support of the National Kidney Research Fund and the Lister Institute in funding this work. C.A.B. is a Lister Institute Research Fellow.

REFERENCES

1 Gossler, A., Joyner, A.L., Rossant, J. and Skarnes, W.C. (1989) *Science* **244**, 463–465.
2 Kerr, W.G., Nolan, G.P., Serafini, A.T. and Herzenberg, L.A. (1989) Cold Spring Harbor Symp. Quant. Biol. **54**, 767–776.
3 Friedrich, G. and Soriano, P. (1991) *Genes Dev.* **5**, 1513–1523.
4 Skarnes, W.C., Auerbach, B.A. and Joyner, A.L. (1992) *Genes Dev.* **6**, 903–918.
5 Wilson, V., Manson, L., Skarnes, W.C. and Beddington, R.S.P. (1995) *Development* **121**, 877–886.
6 Gritz, L. and Davies, J. (1983) *Gene* **25**, 179–188.
7 Boulter, C.A., Aguzzi, A., Williams, R.L., Wagner, E.F., Evans, M.J. and Beddington, R. (1991) *Development* **111**, 357–366.
8 Lupton, S.D., Brunton, L.L., Kalberg, V.A. and Overell, R.W. (1991) *Mol. Cell. Biol.* **11**, 3374–3378.
9 Robertson, E.J. (1987) In Robertson, E.J. (ed.), *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*. IRL Press, Oxford, pp. 71–112.