Mol. Cells, Vol. 11, No. 2, pp. 192-197

Molecules and Cells

Generation of Fusion Genes Carrying Drug Resistance, Green Fluorescent Protein, and Herpes Simplex Virus Thymidine Kinase Genes in a Single Cistron

Seung-Chul Oh[1], Seon-Young Nam[1], Hee-Choong Kwon[2], Chang-Min Kim[2], Jeong-Sun Seo[3], Rho Hyun Seong[4], Young-Ju Jang[5], Yong-Hoon Chung[1], and Hee-Yong Chung[1,*]

[1] *Department of Microbiology, Hanyang University College of Medicine, Seoul 133-791, Korea;*
[2] *Laboratory of Molecular Oncology, Korea Cancer Center Hospital, Seoul 139-706, Korea;*
[3] *Macrogen Incorporation, Seoul 100-799, Korea;*
[4] *Institution of Molecular Genetics, Seoul National University, Seoul 151-742, Korea;*
[5] *Institution of Medical Sciences, School of Medicine, Ajou University, Suwon 442-749, Korea.*

(Received October 20, 2000; Accepted November 21, 2000)

We generated new fusion genes carrying positive- and negative-selection markers, and a reporter gene in a single reading frame. The new genes were constructed by sequentially linking the coding sequences of drug-resistance genes (*hygro,* or *puro*), a green fluorescence protein (GFP) gene (*gfp*), and the thymidine kinase gene (*tk*). The new synthetic genes (*hygro/gfp/tk* and *puro/gfp/tk*) were inserted into retroviral vectors to test their usefulness as selective markers and reporters. The genes were functional in a positive selection in the presence of hygromycin (*hygro/gfp/tk*) or puromycin (*puro/gfp/tk*). In addition, cells expressing the new fusion genes were clearly identifiable by their green fluorescence emitted from GFP. At the same time, these cells were sensitive to a gancyclovir treatment, allowing efficient removal of the transduced cells. The presently described synthetic genes will be valuable tools in both gene therapy and basic gene transfer studies, where positive selection of the transduced cells, monitoring gene expression, and negative selection of the transduced cells are simultaneously required.

Keywords: Reporter; Retroviral Vector; Triple Fusion Genes.

Introduction

Viral vectors are the primary vehicles for gene delivery in human gene therapy and their ability to deliver genes into a broad range of tissues and somatic cells makes them useful systems for this purpose (Mulligan, 1993). The usefulness of the viral vector system is further enhanced by the development of the various marker genes, and vectors with an internal ribosome entry site (IRES) and/or internal promoters (Kang *et al*., 1997; Kim *et al*., 1997; Miller, 1997). The commonly used marker genes include the β-galactosidase (*LacZ*) gene, and the neomycin phosphotransferase gene (*neo*) as a reporter, or as a positive selection maker. The *Herpes simplex* thymidine kinase gene (*tk*) has been used as a negative selection marker in various gene therapy protocols. In addition, various combinations of reporter genes have been used in synthesizing genetically engineered fusion genes with dual functions, such as *LacZ-neo* (Abram *et al*., 1997), *neo-tk* (Schwartz *et al*., 1991), and *LacZ-tk* (Marini *et al*., 1995). Development of these fusion reporter genes has increased the versatility of the vector system in a significant way without compromising the integrity of the vector structure itself.

Recently, a green fluorescent protein (GFP) has been widely used as a reporter molecule for the study of protein localization and gene expression. GFP, which is isolated from the jellyfish *Aequorea Victoria* (Chalfie *et al*., 1994), is comprised of 238 amino acids with a molecular mass

* To whom correspondence should be addressed.
Tel: 82-2-2290-0643; Fax: 82-2-2290-0643
E-mail: hychung@email.hanyang.ac.kr

Abbreviations: GCV, gancyclovir; *gfp*, green fluorescence protein gene; *hygro*, hygromycin phosphotransferase gene; *tk*, thymidine kinase gene.

of 27 kDa. The chromophore responsible for the light emission is a hexa-peptide within the protein among which serine-65, tyrosine-66, and glycine-67 play key roles by auto-cyclization (Cormack *et al*., 1996; Yang *et al*., 1996). The coding sequence of the GFP (*gfp*) has been linked in frame to the drug-selective markers, such as *neo* (*neo-gfp*) (Karreman, 1998), or the hygromycin phosphotransferase gene (*hygro-gfp*) (Lybarger *et al*., 1996) to produce fusion proteins displaying both fluorescence and drug-resistance. The *gfp* was also linked to *tk* to produce a fusion gene (*gfp-tk*) displaying both fluorescence and drug-sensitivity (Loimas *et al*., 1998).

Since GFP has a very stable chromophore, we hypothesized that triple fusion genes with positive- and negative-selection markers and *gfp* may maintain the functional integrity of all three of their components. To test this hypothesis, we constructed retroviral vectors expressing triple fusion genes (*hygro/gfp/tk* or *puro/gfp/tk*), and showed efficient selection (both positive- and negative-) and identification (by fluorescence) of the transduced cells *in vitro*. In addition, we showed that the expression of the newly synthesized fusion genes did not compromise the viral titers in a retrovirus packaging cell line.

We expect that the synthetic genes described herein will serve as valuable tools in both gene therapy, and basic gene transfer studies when positive selection, monitoring gene expression, and negative selection of the transduced cells are simultaneously required.

Materials and Methods

Synthesis of Fusion Genes and Retroviral Vectors Three retroviral vectors, MFG.*tk*, MFG.*gfp*, and MFG.*hygro/gfp* (Fig. 1) were first constructed. MFG.*tk* (Fig. 1A) was constructed by deleting the IRES-neo cassette from MOTEN (obtained from M. Sadelain, Memorial Sloan-Kettering Cancer Center) by BamH1 digestion and self-ligation. MFG.*gfp* (Fig. 1B) was constructed by replacing the heat shock protein 70 (*HSP70*), coding sequence of MFG.*HSP70* (Kwak *et al*., 1998), with the *gfp* coding sequence (digested with Nco-1 and *Bam*H1) obtained by PCR amplification of the pEGFP-N1 plasmid (Clontech Laboratories, USA) by using the following primers: 5'-aaatccATGGTGAGCAAGGGCGAGG-3' (sense orientation) and 5'-aaatggatcctaTTACTTGTACAGCTC-GTCCATG-3' (anti-sense orientation). Both the retroviral backbone and the PCR fragment were digested with a *Nco*1 and *Bam*H1 restriction enzyme before ligation. By a similar ligation, MFG.*gfp.IRES.puro* and MFG.*tk.IRES.puro* were constructed by replacing the *HSP70* sequence of MFG.*HSP70.IRES.puro* with *gfp and tk* respectively (Fig. 1C). MFG.*hygro/gfp* (Fig. 1D) was constructed by performing three parts ligation with the following fragments: i) *Nhe*1 to *Sal*1 fragment (1,825 bp) from pHygro-egfp (Clontech Laboratories, Ltd.), which contained a coding sequence of the *hygro/gfp* fusion gene; ii) BglII to *Xba*1 fragment (359 bp) of the MFG retroviral backbone; (iii) BglII and *Xho*1 digested MFG.I-A^{b} plasmid (obtained from M. Sadelain, Memorial Sloan-Kettering Cancer Center), which contained pUC-19 and the majority of a retroviral backbone sequence, including

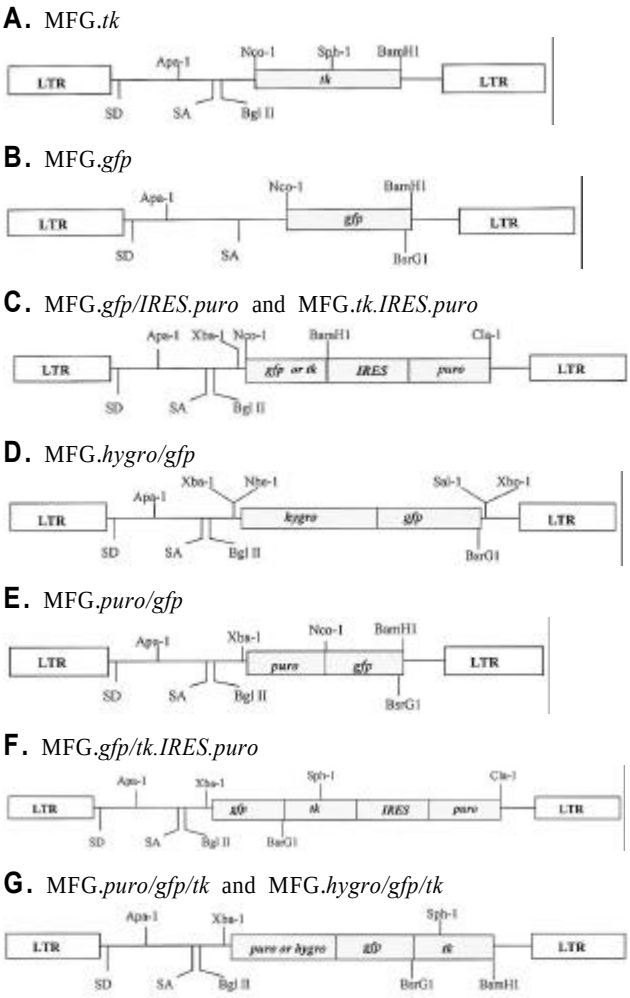

Fig. 1. Schematic illustration of the retroviral vectors used in this study. Details of cloning procedures are illustrated in **Materials and Methods**.

LTRs and a packaging signal (5.1 kb). For MFG.*puro/gfp* (Fig. 1E), the puromycin-resistance gene (*puro*) was amplified from pJWpuro (Morgenstern and Land, 1990) with the following primers: 5'-AAAATCTAGACTGCCATGACCGAGTACAAGC-CCACG-3' (sense) and 5'-AAAACCATGGCACCGGGCTTGC-GGGTCA-3' (anti-sense). The sense primer contained a *Xba*1 restriction enzyme site (TCTAGA) in upstream of the start codon of *puro*. The anti-sense primer contained a *Nco*1 site (CCATGG) after the last codon of the *puroI,* so that the digested PCR product can be linked directly to the *Nco*1 site of *gfpI,* which contained the start codon, ATG. The PCR product was digested with *Xba*1 and *Nco*1. The digested fragment (606 bp) was used for three parts ligation with the following DNA fragments: i) retroviral backbone of MFG.*gfp* opened with *Apa*1 and *Nco*1 (6.3 kb); ii) *Apa*1 to *Xba*1 fragment (828 bp) of retroviral backbone. For MFG.*gfp/tk.IRES.puro* (Fig. 1F), the *tk* gene was first amplified by PCR with following primers: 5'-TTTATGTACAAGATG-GCTTCGTACCCCGGCCATCA-3' (sense) and 5'-TTTAAGAT-CTTCAGTTAGCCTCCCCCATCTC-3' (anti-sense). The sense TK primer contained the *Bsr*G1 restriction enzyme site (which is located close to the end of the *gfp* gene manufactured by Clontech), and the remainder (2 bp) of the *gfp* sequence upstream

of the 5' coding sequence of *tk*. The anti-sense TK primer contained a *Bgl*II site downstream of the stop codon of *tk*. The PCR product was digested with *Bsr*G1 and *Sph*1. The resulting fragment of *tk* (393 bp), and the *Bgl*II to *Bsr*G1 fragment (1.08 kb) from MFG.*gfpI,* were cloned into the retroviral backbone of the MFG.*tk.IRES.puro* (6.74 kb) that was opened with *Bgl*II and *Sph*1 (7.94 kb).

To synthesize MFG.*hygro/gfp/tk* (Fig. 1G), the PCR product of *tk* digested with *Bsr*G1 and *Sph*1 (393 bp), together with the *Bgl*II to *Bsr*G1 fragment (2.15 kb) of MFG.*hygro/gfp*, were cloned into the retroviral backbone of the MFG.*tk* (6.74 kb) that was opened with *Bgl*II and *Sph*1. MFG.*puro/gfp/tk* (Fig. 1G) was constructed by three parts ligation with the following DNA fragments : i) *Bgl*II to *Bsr*G1 fragment (1.68 kb) of MFG.*puro/gfpI,* which contained part of the retroviral backbone and most of the *puro/gfp* coding sequence; ii) *Bsr*G1 to *Sph*1 fragment of MFG.*hygro/egfp/tk* (393 bp, see Fig; 1g); iii) retroviral backbone of MFG.*tk* (6.74 kb) that was opened with *Bgl*II and *Sph*1.

Retrovirus production and virus infection To produce a retrovirus producing cell line, the retroviral vector plasmids were introduced into the PA317 packaging cell line by the calcium phosphate precipitation method (Gorman *et al.*, 1983). Two days later, the drug-resistant cells were selected in a selective medium containing either 2 μg/ml of puromycin (Sigma, USA), or 200 μg/ml of hygromycin B (Sigma, USA). The viral supernatants were harvested from the overnight cultures of the mixed population of the drug-resistant PA317 cells at semi-confluency. The viral infection was performed by incubating NIH3T3 cells with the viral supernatant for 4 h in the presence of 8 μg/ml polybrene (Sigma, USA). Forty-eight hrs after the infection, the cells were placed in a selective medium containing either puromycin or hygromycin, depending on the selective markers included in the retrovirus. After 10 d of selection, the drug-resistant cells were harvested and stored for further analysis.

Flow cytometry The expression of GFP was analyzed by the fluorescence intensity of the transduced NIH3T3 cells, and the data were collected on the flow-cytometer (FACSort, Becton Dickinson, San Jose, CA) at an excitation wavelength of 488 nm, then analyzed with the Cell Quest software (Becton Dickinson).

Drug-sensitivity test with gancyclovir (GCV) The virus-transduced NIH3T3 cells were seeded in 6 well culture plates at 5×10^4 cells/well density, and were treated with different concentrations of GCV for 5 d. The live cells were counted after tryphan blue staining. The relative fraction of surviving cells were calculated and plotted against the GCV concentration.

Immunoblot analysis Cell were harvested, washed in PBS and transferred to microcentrifuge tubes. The pellets were suspended in 100 μl of a lysis buffer [10 mM tris (pH 8.0), 60 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.1 % Nonidet P-40 (NP-40), 1 mM PMSF] and incubated on ice for 5 min. An SDS-PAGE analysis was performed using a Hoefer gel apparatus, and cellular extracts were resolved by 10% SDS-PAGE and electrophoretically transferred to nitrocellulose paper. Subsequently, the nitrocellulose paper was incubated with a rabbit anti-GFP polyclonal antibody (Clontech Inc., USA). Horseradish peroxidase conjugated anti-rabbit IgG (Santa Cruz Biotechnology, USA) was used as the secondary antibody, and the immunoreactive protein was visualized by the chemiluminescence (ECL, Amersham, UK).

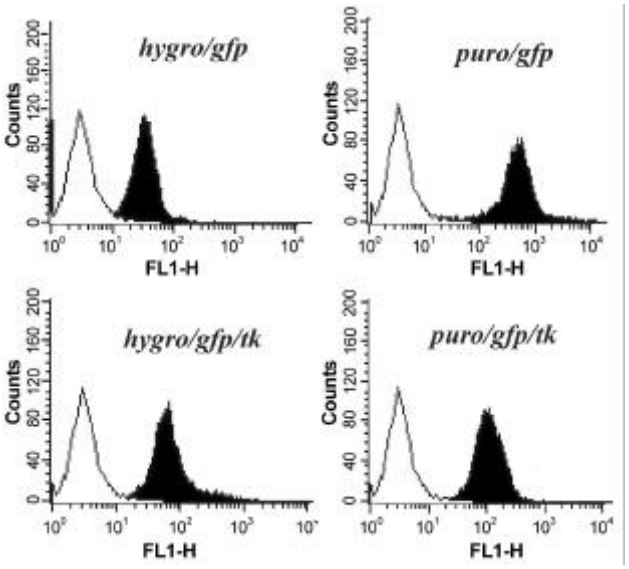

Fig. 2. Flow cytometric analysis of the NIH3T3 cells transduced with various retroviral constructs. NIH3T3 cells transduced with various retroviral vectors were selected with hygromycin B or puromycin. For flow cytometric analysis, the fluorescence was detected in FL1 channel after excitation at 488 nm (black peak). The parental NIH3T3 cells were used as negative controls (gray peak).

Results

Triple fusion gene construction As shown in Fig. 1G, we constructed two triple fusion genes, hygro/gfp/tk and puro/gfp/tk. The gfp gene was positioned at the center, because the other two genes encode enzymes that may need more molecular freedom for appropriate protein folding. The complete coding sequence of each component was linked precisely in frame creating a single uninterrupted reading frame. The sequence was confirmed by DNA sequencing. The synthetic fusion genes were then cloned in a MFG retroviral vector to examine their usefulness in monitoring gene expression, as well as in positive and negative selection.

Functional integrity of the new synthetic fusion genes The functional integrity of the new synthetic genes was tested by using NIH3T3 cells that was transduced with the retroviral vectors encoding the new synthetic genes. The retroviral vector sequence was transferred into the PA317 packaging cell line by DNA transfection, and the cells were selected with either hygromycin or puromycin, depending on the vectors transferred. The drug-resistant retrovirus-producing cells formed discrete colonies in the presence of the selective drugs. These results show that the positive selection marker genes, hygro and puro, are functionally intact as components of triple fusion genes. Fluorescence emitted by the green fluorescence protein in the NIH3T3 cells infected with various retroviral constructs was measured (Fig. 2), and the fluorescence emitted by the new

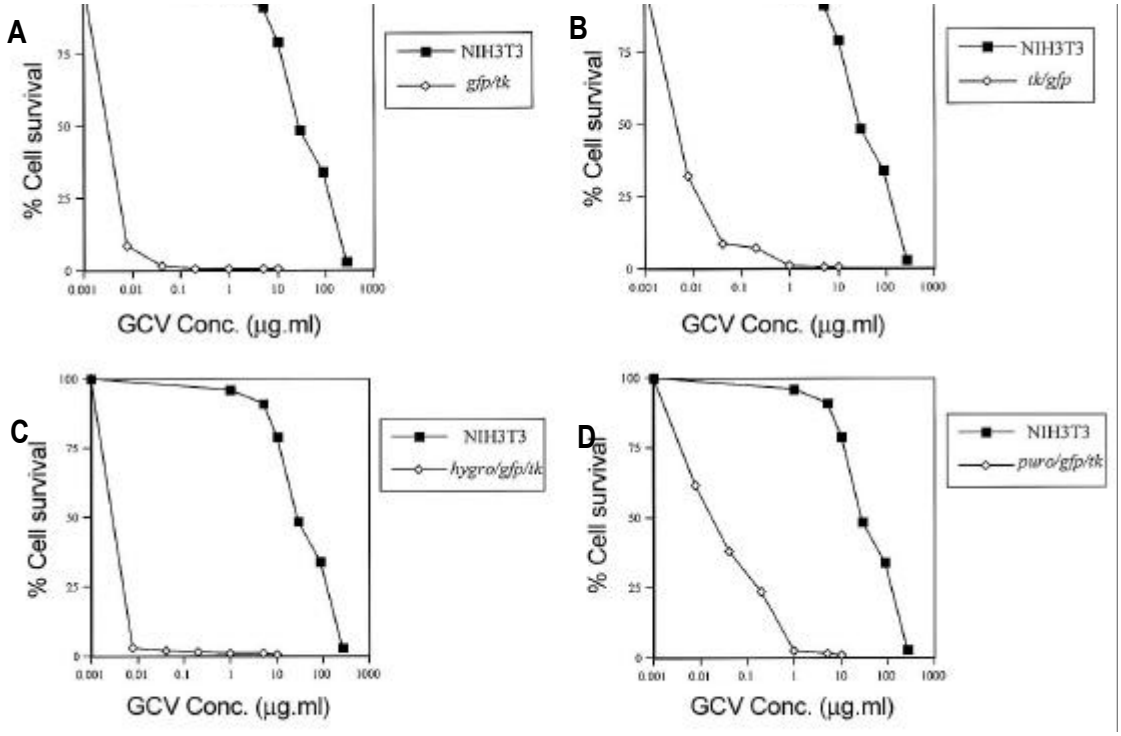

Fig. 3. GCV sensitivity test of the NIH3T3 cells transduced with various retroviral constructs. NIH3T3 cells transduced with various retroviral vectors were selected with hygromycin B or puromycin. For the GCV sensitivity test, the NIH3T3 cells transduced with various retroviral constructs were incubated for 5 d in the presence of serially diluted GCV. At the end of the incubation, live cells were counted after tryphan-blue staining.

triple fusion proteins, as well as the double fusion proteins, was strong enough for efficient detection of transduced cells by a flow-cytometric analysis. We also tested whether the HSV-tk gene (*tk*) in the triple fusion genes was functional. For this, the NIH3T3 cells were transduced with various retroviral vectors, and the transduced cells were tested for their sensitivity to GCV treatment. As shown in Fig. 3, cells expressing the triple fusion genes, as well as those expressing double fusion genes, were sensitive to the GCV treatment. Next, we tested whether the expected sizes of the fusion proteins were synthesized in the transduced NIH 3T3 cells. As shown in Fig. 4, the synthetic genes that we synthesized were expressing the expected sizes of the fusion proteins.

Effects on retrovirus production by the new synthetic genes To test whether or not the new synthetic genes affected retrovirus titers as a component of the retroviral vector, we measured titers of the retrovirus producing cells, after DNA transfection of various retroviral constructs into the PA317 packaging cell line. For a accurate comparison, whole populations of drug-selected packaging cells, instead of single clones, were used to harvest retroviruses. As shown in Table 1, there was no significant difference between the retrovirus titers of the different virus-

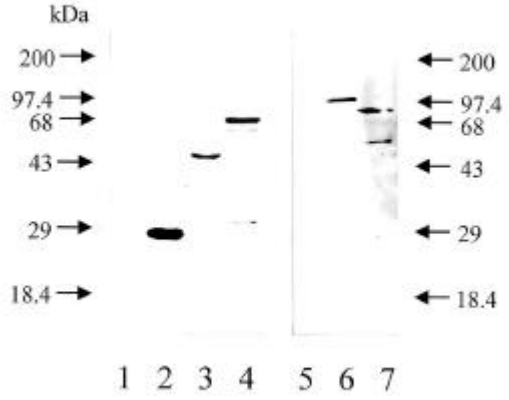

Fig. 4. Immunoblot of fusion proteins. Cellular extracts from the NIH3T3 cells transduced with corresponding retroviral vectors were analyzed for the presence of fusion proteins of appropriate sizes. The GFP-containing proteins were visualized with anti-GFP antibodies and anti-rabbit secondary antibodies. Each lane represents cellular extracts obtained from the NIH3T3 cells transduced with: lane 1, No transduction (parental NIH3T3 cells); lane 2, MFG.*gfp.IRES.puro* (26 kDa); lane 3, MFG.*puro/ gfp* (48 kDa); lane 4, MFG.*gfp/tk.IRES.puro* (68 kDa); lane 5, No transduction (parental NIH3T3 cells); lane 6, MFG.*hygro/ gfp/tk* (106 kDa); lane 7, MFG.*puro/gfp/tk* (90 kDa). See Fig.1 for a detailed structure of the vectors. The extra bands of low molecular weight in lanes 4 and 7 are probably degradation products.

Table 1. Retroviral titers of different constructs after DNA transfection into a PA317 packaging cell line.

Constructs	Retroviral Titers ($\times 10^5$)[a]
MFG.*puro/gfp*	2.5 ± 1.0
MFG.*puro/gfp/tk*	3.1 ± 1.1
MFG.*hygro/gfp*	2.6 ± 0.7
MFG.*hygro/gfp/tk*	3.5 ± 0.9
MFG.*LacZ*	3.1 ± 0.8

[a] Averages of 3 independent measurements (mean ± standard deviation).

producing cells. These results clearly show that the new synthetic genes are compatible with the helper-free retroviral vector system.

Discussion

There have been numerous reports on the synthesis of hybrid genes by linking two genes with different functions. Up until now, various combinations of the *E. coli* β-galactosidase gene (*LacZ*), hygro, neo, tk and gfp genes have been used for this purpose, and these genes have been useful as drug-selection markers, as well as reporters of the gene expression in various gene transfer studies, including gene therapy. Furthermore, insertion of internal promoters, and/or multiple IRES, further increased the versatility of the gene transfer vectors. In contrast to these advantages, however, the promoter interference (Soriano *et al*., 1991) and inefficient translation initiation (Martinez-Salas, 1999) have also been reported in using these types of vectors. In this report, we showed that three genes with selective and reporter functions could be expressed as a single cistron, and that all of the components of the fusion genes were functional. To our knowledge, this is the first report on the synthesis of triple fusion genes with three different functions.

The genes presently described are expected to be useful in various basic and clinical gene transfer studies. Since the positive selection, monitor of the gene expression and/or the fate of the transduced cells *in vivo* (or *in vitro*), and negative selection can be performed by the expression of a single gene, the new synthetic genes will significantly alleviate the structural constraints on the vector structure when the expression of multiple genes and markers is simultaneously required. For example, it is now possible to select for the hematopoietic stem cells transduced with the expression vectors, to monitor gene expression, and also to remove such cells at an appropriate time by expressing a single fusion gene in addition to the gene of interest.

Acknowledgments This research was supported in part by grants [G7 Research and Development Grant to J. S. Seo and Molecular Medicine Research Group Program Grant (#MM-02-02-A-04) to H.-Y. Chung] from the Ministry of Science and Technology, Republic of Korea. The authors also wish to acknowledge the financial support of Hanyang University, Korea, made in the program year of 2000.

References

Abram, C. L., Page, M. J., and Edwards, P. A. (1997) A new retroviral vector, CA1, to identify and select for cells expressing an inserted gene *in vitro* and *in vivo*. *Gene* **196**, 187−189.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994) Green fluorescent protein as a marker for gene expression. *Science* **263**, 802−805.

Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996) FACS-optimized mutants of the green fluorescent protein (GFP). *Gene* **173**, 33−38.

Gorman, C., Padmanabahn, R., and Howard, B. H. (1983) High efficiency DNA-mediated transformation of primate cells. *Science* **221**, 551−553.

Karreman, C. (1998) New positive/negative selectable markers for mammalian cells on the basis of Blasticidin deaminase-thymidine kinase fusions. *Nucleic Acids Res.* **26**, 2508−2510.

Kim, S. J., Sadelain, M., Choi, K. H., Kim, H. K., Lee, J. S., and Chung, H. Y. (1997) Tetracycline-mediated suppression of gene expression with a new dicistronic retroviral vector. *Mol. Cells* **7**, 514−520.

Kang, H.-M., Kang, N. G., Kim, D-.G., and Shin, H. S. (1997) Dicistronic tagging of genes active in embryonic stem cells of mice. *Mol. Cells* **7**, 502−508.

Kwak, H.-J., Jun, C. D., Pae, H-.O., Yoo, J. C., Park, Y. C., Choi, B. M., Na, Y. G., Park, R. K., Chung, H. T., Chung, H. Y., Park, W. Y., and Seo, J. S. (1998) The role of inducible 70-kDa heat shock protein in cell cycle control, differentiation, and apoptotic cell death of the human myeloid leukemic HL-60 cells. *Cell. Immunol.* **187**, 1−12.

Loimas, S., Wahlfors, J., and Janne, J. (1998) Herpes simplex virus thymidine kinase-green fluorescent protein fusion gene: new tool for gene transfer studies and gene therapy. *Biotechniques* **24**, 614−618.

Lybarger, L., Dempsey, D., Franek, K. J., and Chervenak, R. (1996) Rapid generation and flow cytometric analysis of stable GFP-expressing cells. *Cytometry* **25**, 211−220.

Marini, Fc. 3rd., Cannon, J. P., Belmont, J. W., Shillitoe, E. J., and Lapeyre, J. N. (1995) *In vivo* marking of spontaneous or vaccine-induced fibrosarcomas in the domestic house cat, using an adenoviral vector containing a bifunctional fusion protein, GAL-TEK. *Hum Gene Ther.* **6**, 1215−1223.

Martinez-Salas, E. (1999) Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.* **10**, 458−464.

Miller, A. D. (1997) Development and applications of retroviral vectors. in *Retroviruses*, Coffin, J. M., Hughes, S. H., and Varmus, H. E. (eds.), pp. 437−474, Cold Spring Harbor Laboratory Press.

Mogenstern, J. P. and Land, H. (1990) A series of mammalian expression vectors and characterization of their expression of a reporter gene in stably and transiently transfected cells. *Nucleic Acids Res.* **18**, 1068.

Mulligan, R. C. (1993) The basic science of gene therapy. *Science* **260**, 926−932.

Schwartz, F., Maeda, N., Smithies, O., Hickey, R., Edelmann,

W., Skoultchi, A., and Kucherlapati, R. (1991) A dominant positive and negative selectable gene for use in mammalian cells. *Proc. Natl. Acad. Sci. USA* **88**, 10416–10420.

Soriano, P., Friedrich, G., and Lawinger, P. (1991) Promoter interactions in retrovirus vectors introduced into fibroblasts and embryonic stem cells. *J. Virol*. **65**, 2314–2319.

Yang, T. T., Cheng, L., and Kain, S. R. (1996) Optimized codon usage and chromophore mutations provide enhanced sensitivity with the green fluorescent protein. *Nucleic Acids Res.* **24**, 4592–4593.