Investigation of the Transcription in C6 Cells Trapped by a Novel Gene Trap Vector of the Convergent Type

Chika KUTSUWADA1,2, Akinori MIYASHITA1, Tamaki SASAKI1,2, Shoji ODANI2 and Ryozo KUWANO1

1Department of Molecular Genetics, Genome Science Branch, Center for Bioresource-based Researches, Brain Research Institute, Niigata University; 2Course of Biosphere Science, Graduate School and Technology, Niigata University, Niigata, Japan

Received December 27 2002; accepted January 20 2003

Summary. The interference of transcription by antisense transcripts is now regarded as one of the regulatory mechanisms of gene expression. The overlapping of genes in opposite strands has recently attracted much attention because of their potential association with human diseases. The regulatory mechanism of the overlapping gene itself is also of considerable interest from the viewpoint of gene expression. We designed a novel gene trap vector (GTV) that can be transcribed bidirectionally. It contained promoters, dual and orientation-specific reporter genes (β-galactosidase and green fluorescent protein), drug-resistant genes, and SV40 polyadenylation signal flanked by loxP sites. C6 cell lines stably integrated with the GTV were established by selection with antibiotics. Expression patterns of sense and antisense strands around the trap vector examined by the two reporter genes and their transcripts varied substantially by cell lines: some of this was attributable to interference by antisense transcription, but a portion seemed to be under the control of unknown mechanisms. These results proved the present GTV to be useful as a model system for the study of bidirectional transcription in gene expression.

Key words—gene trap, convergent transcription, overlap gene.

INTRODUCTION

The gene trap (GT) strategy is a technique that can identify novel genes and their function, mode of expression, and other characteristics by insertion of a GT vector (GTV) containing reporter genes (e.g. β-galactosidase; lacZ) into a genome. This strategy depends on the random integration of a promoterless GTV, followed by its expression only when the vector is inserted into an active transcription unit of cellular genes. The random integration complements the direct approach of gene targeting, since genes can be accidentally disrupted by the random integration of a GTV. The GTV also has a role as a molecular tag for identification of the trapped gene by cloning. Promoter traps are based on the integration of a reporter gene lacking a promoter into the genome and its expression by the endogenous promoter of a trapped gene. The expression pattern and regulation of the trapped gene can be indirectly investigated by the expression of the reporter gene in the trap vector. This allows to elucidate the mechanism of transcriptional regulation in the trapped cell.

We have established a GT mouse line (GT311) by means of the GT strategy. In this GT mouse, the GTV was inserted between two functional genes, Elf4a1 and Fxr2h, which were on opposite strands, with the concomitant deletion of a 35-kb segment containing three other genes, Cd68, Sup115h, and Sox15, the latter two of which were also arranged tail-to-tail with overlapping 3'-ends and transcribed in opposite directions. In the heterozygous mouse, expression of the GTV reporter gene (LacZ) was restricted to the testis and in a mosaic pattern among

Correspondence: Chika Kutsuwada, Department of Molecular Genetics, Genome Science Branch, Center for Bioresource-based Researches, Brain Research Institute Niigata University 1-757 Asahimachi-dori, Niigata 951-8510, Japan

Abbreviations – Cd68, murine macrophilin gene; Elf4a1, eukaryotic protein synthesis initiation factor 4A1 gene; Fxr2h, fragile X-related gene 2 homolog; NLS, nuclear localization signal; pGK, phosphoglycerate kinase 1, SA, splicing acceptor, Sox15, Sry-like HMG box gene 15; Sup115h, suppressor of Loc15 homolog (C. griseus) gene.
the cells (Miyashita et al. unpublished observation). This is quite curious because the promoter trapped gene, Eif4a1, codes an indispensable factor of protein synthesis and should be expressed ubiquitously. Since an RT-PCR analysis detected antisense transcripts from the trap allele, it was supposed that transcripts of the antisense side in the flanking region of the trapped vector had exerted a certain influence on the expression of the reporter gene in the sense strand. Regulatory phenomena in gene expression, which are collectively termed transcriptional interference, are defined broadly as the perturbation of one transcription unit by another. Mechanisms likely include the following: promoter obstruction; production of antisense RNA associated with inhibition of translation (or RNAi); epigenetic mechanisms like methylation or histone modifications; and direct transcription of one gene by the promoter of another. Increasing evidence suggests that double strand RNAs derived from sense and antisense transcripts participate in these regulatory processes in gene expression.

In order to investigate the influence of transcription of the antisense strand on that of the sense strand, we constructed a new GTV that should be transcribed in the sense and antisense directions with a different set of different reporter genes. The Cre/loxC system was also incorporated into the present GTV to allow conditional transgene expression or gene disruption in mice; this can bypass initial embryonic lethal phenotypes and enable access to phenotypes in later stages in combination with inducible Cre recombinase expressing vectors.

The present paper describes the construction of a novel convergent-type GTV and the consequence of its stable transfection into C6 cells as monitored by the expression of the two reporter genes.

MATERIALS AND METHODS

Construction of gene trap vectors

Structures of vectors designed for the present study are summarized in Fig.1. For preliminary experiments to estimate the efficiency of the transfection and validity of the SV40PA and Cre-loxC system, two vectors, pMoSVpuroLuc (#154, Fig. 1A) and pMoSV-Luc (#145, Fig. 1B), were used; these have an SV40 polyadenylation signal (SV40PA) flanked by a pair of loxC between the promoter and reporter gene (luciferase gene, luc). pMoLucA (Fig. 1C) was for the positive control in a luciferase reporter assay.

pMV3 has a structure in which two marker genes (neo, LacZ), a splicing acceptor (SA), internal ribosomal entry site (IRES), and thymidine kinase polyadenylation signal (kPA) are inserted into pSP719 (Pharmacia Biotech) vector (Fig. 1D). A construction of the splicing acceptor, HygEGFP (Clontech), and SV40PA flanked by a pair of loxC recombination targets was inserted into pMV3 in the opposite direction (pMV3VSConv). To this was further inserted a major immediate early promoter from human cytomegalovirus (hCMV) (Fig. 1D, #193, pMV3VSConvCMV) in the direction of the antisense. This gene was digested by SmaI restriction enzyme for stable transfection into C6 cells. A Cre expression vector, which enables selection by puromycin, was also constructed (Fig. 1E).

Cell culture

C6 cells were cultured in Dulbecco's modified Eagle medium (DMEM, Nissui) containing 10% fetal bovine serum (FBS), 2 mM L-glutamine, and 0.2% NaHCO3. C6 glioma cells (rat) transiently transfected by GTV (#154) and control vectors were collected and solubilized with Luc PGC50 (TOYO Ink). The luciferase activity of these extracts were measured using a luminometer. Cells were transferred to a culture medium without 10% FBS in 6-well plates (1×104/well), incubated at 37°C for 24 h, and transfected with four gene trap vectors digested with SmaI using LIPOFECTIN Reagent (LIFE TECHNOLOGIES Inc.). After incubation for 5 h at 37°C, the cells were incubated for 24 h in a culture medium. After 24 h, selection was performed in a culture medium containing neomycin (600 μg/ml) and hygromycin (350 μg/ml) for ten days. The medium containing these antibiotics was changed every two days. Colonies were picked up for cloning and cultured in 24-well plates using limiting dilution, and then these cells were transferred to other 24-well plates.

Colony forming ability

Cells trapped by GTV, selected for 14 days with G418 and hygromycin, were fixed by 4% paraformaldehyde, and then stained by Mayer's hematoxylin solution at room temperature. After a washing with PBS (-), the number of colonies formed was counted.

Cre-mediated excision

The p2InsCARTpuro (#199) vector was designed to work in the Cre/loxC system in the gene trap vectors which we constructed, containing an hCMV promoter, polyadenylation signal (PA), promoter and
(A) \( \text{pMoSV}_{\text{puroLuc}} (\# 154) \)

\[
\text{MMLV} \xrightarrow{\text{SV40PA}} \text{PGK}_{\text{puro}} \xrightarrow{\text{loxP}} \text{Luc}
\]

(B) \( \text{pMoSVLuc} (\# 145) \)

\[
\text{MMLV} \xrightarrow{\text{SV40PA}} \text{loxp} \xrightarrow{\text{loxP}} \text{Luc}
\]

(C) \( \text{pMoLucA} \)

\[
\text{MMLV} \xrightarrow{\text{Luc}} \text{PA}
\]

(D) \( \text{pMV3VSC} \text{ConvCMV} (\# 193) \)

\[
\text{SAneo} \xrightarrow{\text{RES}} \text{lacZ} \xrightarrow{\text{tAPA}} \text{NLS}
\]

(E) \( \text{p2InsCART7puro} (\# 199) \)

\[
\text{insulator} \xrightarrow{\text{CMVactinNLSCrePA}} \text{PGK}_{\text{puro}} \xrightarrow{\text{PA}} \text{insulator}
\]

Fig. 1. Construction of GTVs. (A) \( \text{pMoSV}_{\text{puroLuc}} (\# 154) \), SV40PA is placed between \( \text{loxP} \) sites upstream of the luciferase reporter gene (Luc). (B) \( \text{pMoSVLuc} (\# 145) \), a construct derived from \# 154 excluding PGKpuro. (C) \( \text{pMoLucA} \), a control vector of (A) and (B). These vectors were designed for preliminary experiments. (D) \( \text{pMV3VSC} \text{ConvCMV} (\# 193) \), a new GTV that carries a convergent construction, which has an \( \text{hCMV} \) promoter upstream of the antisense side. (E) \( \text{p2InsCART7puro} (\# 199) \), Cre recombination vector for Cre/loxP system.

Puromycin resistant gene (Fig. 1E). C6 cells stably transformed by the gene trap vectors were transfected to 6-well plates (1 \( \times 10^5 \)/well), incubated for 24 h at 37°C, and transected transiently with \( \text{p2InsCART7puro} \) using LIPOFECTIN Reagent (LIFE TECHNOLOGIES Inc.) at 37°C for 5 h. After lipofection, the medium for lipofection was changed to a DMEM culture medium and incubated for 24 h. The medium was changed every two days\(^9\).

X-gal staining

Stable GTV-transformant cell lines were examined for \( \beta \)-galactosidase activity before and after transient transfection with the Cre-expressing vector (\( \text{p2InsCART7puro} \)). The cells were washed twice by PBS (--), and fixed for 10 min in 4% paraformaldehyde at room temperature. Fixed cells were then stained for 5 h in a solution containing 1 mM \( \text{MgCl}_2 \), 5 mM \( \text{K}_4 [\text{Fe(CN)}_6]_2 \), 5 mM \( \text{K}_3 [\text{Fe(CN)}_6]_3 \), and 0.1% X-gal at 37°C. After staining, cells were washed in PBS (--), and observed with a microscope\(^{10}\).

Enhanced green fluorescent protein

The expression of GTV carrying a GFP reporter gene was observed for intact cells under a fluorescence microscope\(^{16,18,20}\).

Southern analysis

Genomic DNAs were extracted from cells, and were
digested with *EcoR*I and separated on a 1.0% agarose gel, transferred to a nitrocellulose filter, and then hybridized with a 32P-labeled probe. A 247bp SV40PA fragment (*Bam*HI/*Pst*I treated) was used for the probe.

**RT-PCR**

cDNAs were prepared with total RNA extracted from cells by ISOGEN (Wako Pure Chemicals). The primers used for the first strand cDNA synthesis corresponded to parts of the vector sequence: 5'-ctacacagctcatggctccag-3', 5'-tatattgagcagcagggg-3', 5'-catcgcagcatcaccagcag-3' and 5'-cgatactgacggctccag-3', which were named hygc-1, hygc-2, lacZc-1, lacZc-2, respectively. PCR primers were 5'-gcacattcgaattgcggcgc-3', 5'-acgcattcgaattgcggcgc-3', 5'-gtttgacattaaccgccagcagtgcag-3', and 5'-gtttgacattaaccgccagcagtgcag-3', named hyg-3, hyg-4, lacZ-1, lacZ-2, respectively. Strand-specific reverse transcription-PCR (RT-PCR) was performed under the following conditions: 93°C for 1 min, 60°C for 1 min, and 72°C for 2 min for 30 cycles in a 10 μl solution of 10 mM Tris-Cl (pH 8.3), 50 mM KCl, 1.5 mM MgCl2, 2.5 mM each of dNTPs, 10 μM each of the primers and 0.4 U of *Taq* polymerase (TaKaRa).

**RESULTS**

**Transient expression of reporter genes**

First, the efficiency of SV40PA and the operation of the Cre/loxP system in the C6 cells were verified using control vectors. C6 cells were transfected with the pMoSV*pro*Luc vector (#154, Fig. 1A), composed of a MMLV (Mo1ony murine leukemia virus) promoter, SV40PA flanked by *lox*P sites, and a luciferase reporter gene. Luciferase activity in the cell extracts was measured before and after transfection with a Cre-expressing vector (#199, Fig. 1C). As the construct #154 contains a polyadenylation signal (SV40PA) between *lox*P and upstream of the reporter gene, the luciferase activity of the cells transfected by this vector alone was expected to be low, and increase when the cells were co-transfected with the Cre-expressing vector, thereby excising SV40PA out by site-specific recombination. The observed luciferase activities of the extracts from the transient transfectants with vectors #145 and #154 or a stable transfectant with #154 were indeed very low and were less than 0.2% of those with the control vector without SV40PA (pMoLucA) (Table 1). The results indicate that the promoter activity of the #145 and #154 did not enhance the expression of downstream genes due to the presence of SV40PA. The luciferase activity of the #154 stable transfectant increased 50-fold when the cells were transfected transiently by

---

**Table 1. Luciferase activity of cells trapped by control vectors**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Luc activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>C6</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>#145</td>
<td>3</td>
<td>1.39 ± 0.39</td>
</tr>
<tr>
<td>pMoLucA</td>
<td>3</td>
<td>1996 ± 404</td>
</tr>
<tr>
<td>#154</td>
<td>3</td>
<td>4.4 ± 1.02</td>
</tr>
<tr>
<td>#154 stable</td>
<td>3</td>
<td>0.26 ± 0.19</td>
</tr>
<tr>
<td>#154 stable-#199</td>
<td>3</td>
<td>13.02 ± 1.99</td>
</tr>
</tbody>
</table>
The Transcription in C6 Cells by a Convergent GTV

Fig. 3. Hematoxilin staining of GTV #193 stable transformants which had been selected by antibiotics for 14 days. (A) Neomycin (G418), (B) hygromycin, (C) G418 and hygromycin.

the Cre vector (Fig. 2). This suggests that the promoter activity of cells trapped by control vectors was restored by the Cre recombinase-loxP site-mediated recombination.

**Colony forming ability**

New gene trap vectors (GTVs) were designed to have promoters at both 5'-ends of the sense and antisense strands, to be convergently transcribed from both sides. The vectors had lacZ and neo at the sense side, and HygEGFP at the antisense side for reporter genes. A polyadenylation signal sequence of SV40 flanked by loxP was also inserted for these GTV to be manipulated with the Cre/loxP system. The Cre-expressing vector was designed to select cells by puromycin to report whether the Cre/loxP system was effective.

**Gene expression in cells trapped by #193**

C6 cells stably transfected with the GTV were selected both by hygromycin and G418 or with either of them. After selection, the colonies formed were stained by hematoxilin, and the numbers of colonies were counted to estimate the colony forming ability in order to deduce the efficiency of the integration of the GTV. Though colonies selected with antibiotics of the promoter containing side were for more numerous than those of colonies selected with antibiotics of the promoterless sides, small numbers of colonies were always present when selected with the drug of the promoterless side or drugs of both sides (Fig. 3). The mode of expression of the cells was examined by X-gal staining and the fluorescence of GFP. Cells integrated with #193, which carries a promoter in the antisense side, were selected by hygromycin, and surviving colonies were stained by
C. Kutsuwada et al.

**Phase**

**GFP**

**X-gal stain**

**Fig. 4.** Fluorescence of GFP and X-gal staining of #193 stable transformant. (A) GFP fluorescence of #193 transient transformants observed by fluorescence microscopy. The right photographs are phase contrast microscopic images of the same fields. (B) X-gal staining of #193 stable transformant at day 14 of selection.

X-gal. They were stained positively by X-gal, though very faintly, and showed a fluorescence of GFP as well (Fig. 4). Next, the cells trapped by #193 were transiently transfected by the Cre vector (#199) to investigate the expression pattern of the cells when SV40PA was removed. The transformants further transfected with the Cre vector were negative in GFP fluorescence, but were positive by X-gal staining (Fig. 4A and B).

**Establishment of cell lines**

GTV integrated cells were selected by G418 and hygromycin for 14 days, and five colonies were selected for cloning by limiting dilution, from which cell lines #193-1 and #193-2 were established and characterized as described below.

**Analysis of integration of GTV by Southern blotting**

The presence of the trapped gene in the genome of cells was confirmed by PCR and Southern blot analysis. In PCR analysis, primers were designed to amplify a sequence (917 bp) including SV40PA (260 bp) and loxP sites (Fig. 5A). Fig. 5B shows the results of PCR for the two cell lines, #193-1 and #193-2, before and after transfection with the Cre vector. The expected 917bp segment of GTV #193 was actually present in the genome of the two cell lines (Fig. 5A, lanes 1 and 3). In addition, the Cre/loxP system appeared to have operated effectively within these
Analysis of expression of the convergent GTV by RT-PCR

Expression of the GTV was analyzed by RT-PCR. Since the present GTV may be transcribed in both directions, first strand cDNAs were synthesized from total RNA using both sense and antisense strand-specific RT primers set at 5' ends of the sense and antisense strands of lacZ and HygEGFP genes. (Fig. 6A). First strand cDNAs were further amplified by PCR with primer sets directed at the lacZ and HygEGFP regions (Fig. 6A). The results indicate that, in the cell line #193-1 (Fig. 6B, left panel), bands of PCR-products corresponding to the lacZ region were much more intense in lacZc-2-primed cDNA than in lacZc-1-primed cDNA, suggesting that the lacZ region was transcribed much more strongly in
Fig. 6. RT-PCR analysis of trapped cells. (A) Position of primers for RT-PCR. Thick, long arrows indicate orientation-specific cDNA synthesis of transcripts by reverse transcriptase. Short, thin arrows are primers for locus-specific cDNA amplification by PCR. (B) Analysis of RT-PCR products by electrophoresis using 1% agarose gel. RT-PCR product of GAPDH was observed using random primed cDNA as a template.
the sense direction than the antisense direction. In the case of HygEGFP transcription, an intense signal was detected for the PCR product derived from the hygc-2-primed (i.e. sense direction of HygEGFP gene) cDNA pool but not from the hygc-1-primed cDNAs. Similar to this was the case for #193-2 cell line (Fig. 6B, right), although a weak expression was seen in cDNA primed by hygc-1. When Cre was co-expressed, the intensity of PCR products was decreased in lacZc-2 cDNA and hygc-2 cDNA of #193-1, while the PCR product of the latter was increased in #193-2 cells.

**DISCUSSION**

One of the purposes of the present study was to develop a novel gene trap vector which can be transcribed convergently from both strands, and their examine its effect on gene expression when it is integrated into the cellular genome. The idea came from the following unexpected observations. In a previously established GT mouse strain, GT3-11, integration of the GTV was accompanied by the partial deletion of *Eif4a1* and *Fxr2h* genes and the complete elimination of three other genes, *Cd68*, *Sup115h*, and *Sox15*, that had been clustered between the two genes in the wild type. Despite the absolute requirement of an *Eif4a1* product, an initiation factor in protein synthesis, expression of the GTV reporter gene under the control of an *Eif4a1* promoter in heterozygous mice was strictly confined to the testis and was even not uniform among the testicular cells. One possible explanation for the strange pattern of X-gal expression may be interference by the antisense transcription; because the *Eif4a1* and *Fxr2h* gene is arranged in a tail-to-tail fashion between which the trap vector is located, the antisense transcripts from the trapped allele were actually present as examined by RT-PCR. Furthermore, a detailed analysis of the gene organization in the deleted region revealed that *Sup115h* and *Sox15* had been also arranged in opposite directions so that their 3'-ends overlapped each other in the wild type (before integration of the GTV) mouse genome. Regulation of the expression of this unusually condensed cluster of bidirectionally organized genes may be a very complex one. In order to investigate a possible interference between sense and antisense transcripts in gene expression, we designed a bidirectionally transcribable GTV that could be a model of such a system, and examined phenomena in gene expression when transfeceted in C6 cells. Preliminary experiments using control vectors confirmed that SV40PA and a Cre/loxP recombination system could work effectively in C6 cells as examined by expression of the reporter genes (Fig. 2 and Table 1). Next, the convergent GTV (#193) was transfecteed, and stably integrated cells lines were established by selection with antibiotics. Although cells selected by hygromycin showed higher colony-forming rates than those selected by neomycin (G418), as the hygromycin-resistant gene was under the direct control of an hCMV promoter, the latter also formed substantial numbers of colonies. This result indicates that the present vector had been inserted into active transcription units and effectively operated as a promoter trap vector (Table 2). PCR analysis of genomic DNA with primers set at the SV40PA region confirmed the integration of the trap vector in the genome as well as the efficient removal of SV40PA flanked by loxP sites by Cre recombinase (Fig. 5).

**Table 2. Colony forming ability of cells trapped by GTV**

<table>
<thead>
<tr>
<th>Drug</th>
<th>N</th>
<th>Count</th>
<th>Relative value</th>
</tr>
</thead>
<tbody>
<tr>
<td>C6 G418, Hygromycin</td>
<td>3</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>#193 G418</td>
<td>3</td>
<td>29.7±14</td>
<td>7.4</td>
</tr>
<tr>
<td>Hygromycin</td>
<td>3</td>
<td>48±7</td>
<td>12</td>
</tr>
<tr>
<td>G418, Hygromycin</td>
<td>3</td>
<td>4±4</td>
<td>1</td>
</tr>
</tbody>
</table>

The Transcription in C6 Cells by a Convergent GTV
scripts amplified with PCR-primers lacZ1 and lacZ2 were only slightly increased or remained unchanged by Cre recombination in both cell lines.

In the antisense-direction (antisense direction of antisense strand genes under the control of the hCMV promoter), no transcript was detected for the HygEGFP region in #193-1, and a barely detectable amount was amplified after Cre recombination, while the reverse was the case for #193-2. Similarly, a sense transcript from lacZ region was distinctly present in #193-1 cells but decreased after removal of the polyadenylation signal by recombination. In #193-2 cells, however, no marked difference was seen before or after recombination.

These results suggest that, by artificially introducing an overlapping transcription unit into a genome using the bidirectional gene trap vectors designed here, we can induce a wide variety of transcriptional phenomena in C6 cells, some of which can be interpreted as a result of interference of the transcription of the opposite strands such as RNAi, while others are difficult to explain, suggesting some hitherto unrecognized gene-regulatory phenomena. Detailed analysis of the genomic structure of the trapped sites in the genome of individual cell lines may afford some clues to understanding the present results and novel aspects of gene regulation. In the future, we will perform experiments such as a direct sequence to examine in detail the region where it is trapped. In the future, we will perform experiments such as a direct sequence to examine the region where it is trapped.

Eukaryotic overlapping genes, especially those in an anti-parallel fashion, are very rare but have been increasingly recognized for their relationship to human diseases such as amyotrophic lateral sclerosis, homocystinuria, and Alzheimer’s disease.

The presence of overlapping genes should bring about difficulties in transcription, and it has been further suggested that overlapping gene structures increase the propensity towards rearrangements of genes and/or disruption of transcription regulation. In certain instances, bidirectional transcription should occur and affect the expression of the genes. However, how the transcriptions of both strands interfere with each other and ultimately lead to a disease state, or whether other mechanism underlies the pathogenicity is not known. Our bidirectionally transcribable gene trap vector may also be applicable in an investigation of the pathology of the overlapping genes.

Acknowledgments. This work was supported by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology of Japan (R. K).

REFERENCES

15) Sameshima JH, Wek RC, Hatfield GW: Overlapping transcription and termination of the convergent
The Transcription in C6 Cells by a Convergent GTV


