Cloning and characterization of a novel gene with alternative splicing in murine mesenchymal stem cell line C3H/10T1/2 by gene trap screening

Mingke Wang1,2,3,*; Huiqin Sun1,*; Fan Jiang2; Jing Han1; Feng Ye2; Tao Wang1; Yongping Su1 & Zhongmin Zou2,*

1Institute of Combined Injury, State Key Laboratory of Trauma, Burns and Combined Injury, College of Preventive Medicine, Third Military Medical University, 2Department of Chemical Defense and Toxicology, College of Preventive Medicine, Third Military Medical University, Chongqing 400038, 3Department of Internal Medicine, Navy Haitong Hospital, Fuding 355200, P.R. China

A novel gene, designated mgt-6, containing four splicing variants, was isolated from a gene trap clone library of C3H/10T1/2 cells transfected with retroviral promoterless gene-trap vector, ROSAFARY. The transcript variants were differentially expressed in murine tissues and cell lines and differentially responded to diverse stimuli including TGF-β1 and mitogen-activated protein kinase (MAPK) inhibitors. The mgt-6 gene encoded a protein of 37 or 11 amino acid residues with cytoplasmic distribution. However, when C3H/10T1/2 cells were treated with 5-azacytidine, the protein translocated into cell nucleus as indicated by fused LacZ or C-terminally tagged EGFP. Our preliminary results suggest that further study on the role of mgt-6 gene in cell transformation and differentiation may be of significance. [BMB reports 2010; 43(12): 789-794]

INTRODUCTION

Gene trap is a powerful tool to characterize novel genes and analyze their importance to biological phenomena (1, 2). The approach uses a class of reporter vector, that has been designed as containing a splice acceptor site upstream the reporter gene that includes the β-galactosidase (LacZ) gene, the neomycin resistance gene (neo) or EGFP (3), and a splice donor site downstream the selective gene. When the vector integrates into a transcribed gene, the expression of the reporter gene in the gene trap cassette is under the control of the trapped endogenous promoter. Using this strategy, the spatial and temporal expression patterns of the trapped gene under different conditions can be monitored. The fusion transcripts generated from the endogenous gene and the genes in the gene trap vector can be readily identified by rapid amplification of cDNA ends (RACE) technique (4, 5).

Murine embryo-derived mesenchymal stem cell line, C3H/10T1/2 (6) have the potential to differentiate into a variety of specialized tissue cells such as osteocytes, chondrocytes, adipocytes, smooth muscle cells (SMCs) (7-10) and endothelial cells (11, 12). It provides a unique model for examining the molecular regulation of both the developmental determination of vertebrate stem cell lineages and their subsequent differentiation (13).

In this study, we have raised about one hundred gene trapped clones from C3H/10T1/2 cells transfected with gene trap vector ROSAFARY, including 6 cell clones with positive LacZ staining under normal culture condition. The trapped gene sequence in one of the trapped cell clone 6 was obtained and designated as mgt-6 (murine gene trapped clone 6), and its expression pattern, transcript variants, response to various treatments were preliminarily analyzed.

RESULTS AND DISCUSSION

Establishment of a library of cell clones with integrated gene trap vector

Gene trap screens have been used to obtain sequences from large numbers of randomly trapped genes, and the value of reporter gene is that the initial cytological analysis with microscope gives immediate detailed information about subcellular localization (and hence possible function) in the mammalian cells (14, 15). The gene trap vector, ROSAFARY used in present study, contains a promoter-less LacZ-neo fusion (βgeo) reporter gene with a 5′ splice acceptor preceding the ATG site of the βgeo, and a selective gene driven by a PGK promoter and followed by a 3′ splice donor (16) (Fig. 1A). Hygromycin selection allows the cell colonies with the integrated vector to develop, and the β-gal protein that is fused to the transcript of a trapped gene can be easily detected by histochemical staining. The staining also displays the resulting chimeric protein to a variety of different subcellular compartments and this pre-
Cloning a novel gene in 10T1/2 cells by gene trapping

Mingke Wang, et al.

790

Table 1. Exon-intron information of mgt-6 gene

<table>
<thead>
<tr>
<th>Variants</th>
<th>Exon Size (bp)</th>
<th>3' Splice acceptor</th>
<th>5' Splice donor</th>
<th>Intron size (bp)</th>
<th>CDS size (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcript 1</td>
<td>42</td>
<td>...CCAGTgtgag...</td>
<td>3,018</td>
<td>26</td>
<td>23</td>
</tr>
<tr>
<td>Transcript 2</td>
<td>42</td>
<td>...gtgagTGTGC...</td>
<td>281</td>
<td>24</td>
<td>26</td>
</tr>
<tr>
<td>Transcript 3</td>
<td>42</td>
<td>...tgcagTGTGA...</td>
<td>460</td>
<td>22</td>
<td>28</td>
</tr>
<tr>
<td>Transcript 4</td>
<td>114</td>
<td>...tgcagTGTGA...</td>
<td>608</td>
<td>33</td>
<td>33</td>
</tr>
</tbody>
</table>

Fig. 1. The strategy of gene trap, the formation fused transcripts with gene trapped mgt-6, and the products of translation. (A) The strategy of gene trap. The gene-trap vector ROSAFARY contains a promoter trap module (SAβgeo*pA) and a poly-A trap module (PGKhygSD). After inserting into an intron of an endogenous gene at a permissive site and in the correct orientation and reading frame, the promoter trap module can form fusion transcripts, which is driven by the promoter of the trapped gene. In our case, the 5' transcript fragment of the trapped mgt-6 fused to LacZ transcript, and the 3' transcript fragment of mgt-6 fused to hygromycin transcript in the poly-A trap module. Gene-trapped cell clones can be obtained by hygromycin selection, and the trapped genes can be cloned by 5'RACE or 3'RACE. (B) Patterns of β-galactosidase activity in gene trap clones. The β-galactosidase activity (blue color) was uniformly distributed in cytoplasm of cell in clone 6 (a) and another clone (c), or confined to small area (blue dots) in cytoplasm (b and d). Original magnification ×400. (C) The electrophoretogram of RT-PCR products using the specific primers to endogenous gene and vector-derived SA-LacZ gene. A partial sequences of the RT-PCR products was presented. Lower-case letters represented sequences derived from the gene-trap vector (SA and part of LacZ) and the start codon of LacZ was in the red box. Upper-case letters represented the sequence of 5' exon of the trapped gene mgt-6, and the start codon was in the blue box.

Achieving the full-length cDNA sequence of gene trapped in clone 6 by RACE

To obtain the cDNA sequence of the trapped genes, RACE technique was applied. The sequence of the trapped gene in clone 6 turns to be a novel gene when BLAST algorithm was used to search in GenBank and other available databases. The gene was named as mgt-6 (murine gene trap clone 6). The mgt-6 gene was mapped to mouse chromosome 14, and its 3 exons spanned 4.2 kb (Fig. S1A, Table 1). All the sequences at the exon-intron junctions obey the AG-GT rule of splicing. The gene trap cassette was inserted into the first intron (Fig. 1C) as elucidated by RT-PCR and sequencing.

Identification and analysis for spliced transcripts of mgt-6 in C3H/10T1/2 cells

To verify the RACE-derived cDNA sequence and clone the full-length cDNA of mgt-6, primers were designed to span all of the exons of predicted mgt-6 gene, and RT-PCR was conducted on RNA isolated from C3H/10T1/2 cells. To our surprise, the result demonstrated the existence of alternative splicing in exon 2, which led to 4 splicing variants (Table 1, Fig. S1A). The lengths of the splicing transcripts were 982 bp (transcript 1, GenBank accession no. FJ744746), 803 bp (transcript 2, FJ860514), 655 bp (transcript 3, FJ860513), and 374 bp (transcript 4, FJ748867). Transcript included exon 1, 2 and 3. Transcript 2 included exon 1, beginning and middle parts of exon 2, and exon 3. Transcript 3 cDNA included exon 1, beginning part of exon 2, and exon 3. Transcript 4 cDNA included exon 1 and exon 3.
Although the full-length transcripts and its variants of mgt-6 were obtained by RACE technique, the translation starting site needs to be identified. Since LacZ was correctly expressed in clone 6, it is possible to predict the ORF of mgt-6 based on the ORF of the fused LacZ. Since the precise splice acceptor site of ROSAFARY has not been annotated before, we resolved the sequence by RT-PCR using gene specific primer mgt-6-S and the vector specific primer Rosa-AS, which span the junction site of 5' Mgt-6 and LacZ fusing transcript (Fig. 1C). This sequence provided important information for deducing the ORF site of 5' the vector specific primer Rosa-AS, which span the junction site of 5' Mgt-6 and LacZ fusing transcript (Fig. 1C). This sequence provided important information for deducing the ORF of the correctly expressed LacZ in clone 6. In addition, it also confirmed the gene trap vector insertion and proper splicing with the exon upstream of the inserted cassette. Following the above principle, the coding DNA sequences (CDS) of the four splicing variants of mgt-6 were assumed correspondingly. Three of the splicing variants shared a short CDS (MGT-6S) that encoded 13 amino-acids, while one splicing variant had a long CDS (MGT-6L) that encoded 37 amino-acids. Except the last amino acids, the predicted MGT-6S was the same as 5' part of MGT-6L (Fig. S1B). The amino acid sequences did not show any homology to known motif.

Differential expression of the four transcripts
Since four transcripts of mgt-6 existed, their expression pattern in different murine tissues and cell lines was further studied. As shown in Fig. 2A and C, the expression level of the four splicing variants varied in liver, kidney, spleen, intestine, heart, skeletal muscle, brain, skin and lung. The transcript 4 was the only one expressed in kidney, spleen and intestine, and predominately presented in kidney, heart and skeletal muscle, while the transcript 1 predominantly expressed in liver. All four splicing variants expressed in brain and lung. As shown in Fig. 2B and D, C2C12 predominantly expressed the transcript 1, while all four splicing variants expressed in both C3H/10T1/2 and Lewis cells. The expression of transcript 1, 2 and 4 were more strongly in Lewis cells than that in C3H/10T1/2 cells. The significance of this expression profile remains to be explored.

To get the clue of function of mgt-6 gene, the cells of clone 6 were treated with different stimuli in order to see the changes of mgt-6 expression. The cells showed increased strength of LacZ staining after the treatment of PI3K/AKT inhibitor LY294002 (Fig. 3A-d) or p38/RK inhibitor SB203580 (Fig. 3A-f), while weakened LacZ staining after the treatment of ERK inhibitor PD98059 (Fig. 3A-e) or TGF-β1 (Fig. 3A-b). As shown in Fig. 3C and D, the levels of transcript 1, 2 and 3 in C3H/10T1/2 cells were increased after the treatment of LY294002 and SB203580. The transcript 1, 2, 3 and 4 were all downregulated after TGF-β1 stimulation, while only the transcript 1 and 4 were downregulated by PD98059.

It is known that TGF-β1 induces the smooth muscle cell (SMC) phenotype in C3H/10T1/2 cells, which can be attenuated by LY294002(7), while PD98059 can abrogate bFGF-mediated suppression of TGF-β1-induced SMC gene expression (19), implying that PI3K/AKT activation and ERK inhibition may promote the SMC differentiation. We found that the mgt-6 gene expression was downregulated in C3H/10T1/2 cells when treated with TGF-β1 or PD98059, but was upregulated when treated with LY294002 as indicated by LacZ staining and RT-PCR. When analyzing the differential expression pattern of the four transcripts of mgt-6 by semi-quantitative RT-PCR, transcript 1 which encodes the short MGT-6 protein predominantly contributes to the changes of expression after different treatments. The results suggests that the MGT-6 protein is inversely correlated with TGF-β1-evoked cell signaling and may be involved in the process of SMC differentiation.

Subcellular localization of LacZ- or EGFP-fused MGT-6 protein in C3H/10T1/2 cells
5-azacytidine, a DNA methylation inhibitor, can cause neo-
Cloning a novel gene in 10T1/2 cells by gene trapping
Mingke Wang, et al.

Fig. 3. The differential expression pattern of mgt-6 gene after different treatment. (A) The change of β-galactosidase activity in clone 6 induced by different stimuli. (a) Cells from clone 6 normally expressed mgt-6 in cytoplasm as indicated by the fused LacZ downstream. The LacZ staining signal was increased after treated with LY294002 (d) and SB203580 (f) for 24 hours, but decreased after PD98059 (e) and TGF-β1 (b) treatment. Nuclear translocation could be seen after 5-azacytidine (c). Original magnification x200. (B) The subcellular localization of long (L) and short (S) forms of MGT-6 visualized by tagged EGFP in C3H/10T1/2 cells. The EGFP-tagged MGT-6 displayed cytoplasmic distribution in C3H/10T1/2 cells 48 hours after the transfection of pEGFP-N1-M6L (g) and pEGFP-N1-M6S (h), comparing with pan-cyto distribution in pEGFP-N1 transfected cells (i). The nuclear translocation of MGT-EGFP fusion protein (j, k), but not EGFP itself (l) could be observed 9 days after the transfection and pulse azacytidine treatment. Original magnification x200. (C) RT-PCR analysis of mgt-6 expression in C3H/10T1/2 cells after different treatments. (D) Quantification of RT-PCR results. The densitometry analysis was performed by adjusting the original data with that of GAPDH and subtracting the background. Data are shown as means ± SD (n = 3). (a-d) P < 0.05 vs TGF-β1, LY294002, PD98059, SB203580 respectively.

plastic transformation (20) and differentiation of C3H/10T1/2 cells into functional muscle cells, chondrocytes, and adipocytes (21, 22) in addition to the display of corresponding morphologies (23). Transformed phenotype were associated with hypomethylation of the corresponding cellular DNA sequences and the induced expression of certain endogenous retroviral sequences may be one of a series of critical events during the course of multistage carcinogenesis (24). In our study, cells of clone 6 gave a LacZ staining pattern of pan-cytoplasmic distribution (Fig. 3A-a), but the localization of β-gal activity in cells of clone 6 translocated from the cytoplasm to the cell nucleus after treated with 5-azacytidine for 9 days (Fig. 3A-c), which indicates the protein may function inside the nucleus.

To validate this translocalization of MGT-6 protein after 5-azacytidine treatment, EGFP-tagged expression vectors were constructed and transfected into C3H/10T1/2 cells. It was demonstrated that EGFP-tagged long (pEGFP-N1-M6L) or short (pEGFP-N1-M6S) forms of MGT-6 both were predominantly localized in the cytoplasm (Fig. 3B-g, h) under regular culture condition as the control EGFP did (Fig. 3B-i). However, the EGFP-tagged proteins, not the control EGFP (Fig. 3B-i), translocated to nucleus after treated with 5-azacytidine (Fig. 3B-j, k). In addition, the morphology of cells with nuclear translocation of MGT-6 lost the typical spindle-like morphology as revealed by LacZ staining or fluorescence tracking, which indicates the possible transformation of the cells. P38 MAPK can function as a tumor suppressor and negatively regulate tumorigenesis (25, 26), however p38/RK inhibitor SB203580 upregulated mgt-6 expression (Fig. 3A-f). These results give the clues that the novel gene mgt-6 might participate in the process of DNA hypomethylation and the reactivation of some genes involved in cancer development.

In conclusion, we have isolated a novel gene mgt-6 from gene trapped cell clones. The mgt-6 gene contains four transcripts, and these splicing variants differ in their expression in various murine tissues and cell lines. Moreover, the two encoded proteins of mgt-6 undergo nuclear translocation when treated with 5-azacytidine. In addition, the differential expression pattern of the four transcripts in cells treated with TGF-β1, PI3K/AKT and ERK inhibitor indicates the MGT-6 protein inversely correlated with TGF-β1 induced cell signaling. Our preliminary data suggest that MGT-6 might be involved in cell transformation and differentiation.

MATERIALS AND METHODS

Cell culture and viral infection
C3H/10T1/2, C2C12, Lewis cells (ATCC) were grown in DMEM (HyClone) containing 10% FBS with antibiotics. Viral infection and the establishment of a library of gene trapped cell clones from C3H/10T1/2 cells transfected with retroviral promoterless gene-trap vector ROSAFARY (a gift from Dr. Soriano, P), were performed as previously described (18).
LacZ staining
To scan for clones that had been trapped downstream an active promoter, the gene trapped C3H/10T1/2 cell clones were individually inoculated into 96-well plate with normal culture medium. LacZ expression in the subconfluent cells of each cell clone was revealed with LacZ staining kit (Active Motif, CA, USA), and 6 clones were identified with high level of β-galactosidase (β-gal).

RACE
Total RNA was isolated from gene trapped cell clones using the TRIzol (Invitrogen, USA). The 3'RACE was performed by employing the SMARTM RACE cDNA Amplification Kit (Clontech) following the user manual. All primers used in this study are listed in Table S1 (Supplement data). cDNA synthesis was carried out with the primer 3’ CDS using reverse transcriptase M-MLV(RNase H-) (Takera). Then touch down PCR was used with primers 3’ GSP and UPM in 50 μl reaction volume. The thermal cycling parameters employed were: 94°C for 1 min; 94°C for 30 s, 70°C for 30 s, and 72°C for 3.5 min (×5 cycles); 94°C for 30 s, 68°C for 30 s, and 72°C for 3.5 min (×5 cycles); 94°C for 30 s, 66°C for 30 s, and 72°C for 3.5 min (×27 cycles). The final extension step was achieved at 72°C for 5 min. The 5'RACE was performed by employing the GeneRacerTM kit (Invitrogen) following the manufacturer’s guidelines. The strategy of the kit ensured that only the transcripts with 5’cap were reversely transcribed in order to enhance the efficacy of amplifying the 5’ fragment of the trapped gene. cDNA synthesis was carried out using the primer oligo dT. To obtain 5’ ends, PCR amplification of the cDNA was performed using 5’ GSP and 5’ Primer A with the thermal cycling parameters: 94°C for 2 min; 94°C for 30 s, 72°C for 3 min (×5 cycles); 94°C for 30 s, 70°C for 30 min (×5 cycles); 94°C for 30 s, 62°C for 30 s, 68°C for 3 min (×25 cycles); 68°C for 10 min. For rare transcripts, a second round of nested PCR was performed using 5’ GSP and 5’ NUP with the thermal cycling parameters: 94°C for 2 min; 94°C for 30 s, 62°C for 30 s, 68°C for 3 min (×30 cycles); 68°C for 10 min. Analyze 3 μl 5’ or 3’ RACE products on a 1.2% agarose gel, which was stained with ethidium bromide and photographed under UV light. The expression of housekeeping gene glyceraldehyde-3-phosphate dehydrogenase (GAPDH) served as an internal control in all RT-PCR reactions. To quantify the RT-PCR result, the images of the corresponding gels were analyzed by Image-Pro Plus software (Media Cybernetics Inc, USA). The data were adjusted by the expression of internal control GAPDH.

Different treatment, LacZ staining and RT-PCR assessment
The cells of clone 6 were subjected to the treatment of 5 ng/ml TGF-β1 (Peperotech, Inc.), 30 μM LY294002 (PI3K/AKT inhibitor, Beyotime), 70 μM PD98059 (ERK inhibitor, Beyotime), 50 μM SB203580 (p38/RK inhibitor, Beyotime) for 24 hours, or 10 μM 5-azacytidine (Sigma) for 24 hours before switching to regular culture up to 9 days. The cells were stained every the other day for LacZ localization and intensity. After different treatments, the expression patterns of mgt-6 in C3H/10T1/2 cells were assessed by RT-PCR.

Expression of EGFP tagged mgt-6 in vitro
The predicted long coding sequences (CDS) of mgt-6 was obtained using primers M6L-S and M6L-AS by RT-PCR and sequenced. The double-strand DNA fragment of predicted short CDS was obtained by directly annealing the synthesized complementary oligonucleotides M6S-S and M6S-AS (Invitrogen). The long (M6L) and short (M6S) CDS were subcloned into pEGFP-N1 vector (Clontech) respectively and named as pEGFP-N1-M6L and pEGFP-N1-M6S. C3H/10T1/2 cells were transfected with either pEGFP-N1-M6L or pEGFP-N1-M6S using Lipofectamine 2000 (Invitrogen) and cultured for 48 hours. Then the transfected cells were treated with 10 μM 5-azacytidine for 24 hours before switching to regular culture medium for up to 9 days. The GFP expression and location in cells were examined daily.

Statistical analysis
A densitometry analysis of three independent RT-PCR experiments was conducted by One-way ANOVA test using SPSS.
13.0. P < 0.05 was considered statistically significant.

Acknowledgements
This work was supported by grants from the PLA Medical Research Program of Eleventh Five-year Plan (06H029 and 06H030), Development Program from State Key Laboratory of Trauma, Burns and Combined Injury (SKLZZ200811) and the National Key Basic Research and Development Plan (2005 CB522605), China.

REFERENCES


