

# Improvement of reporter activity by IRES-mediated polycistronic reporter system

Hicham Bouabe<sup>1,\*</sup>, Reinhard Fässler<sup>2</sup> and Jürgen Heesemann<sup>1</sup>

<sup>1</sup>Max von Pettenkofer-Institut für Hygiene und Medizinische Mikrobiologie, Ludwig-Maximilians-Universität, 80336 Munich and <sup>2</sup>Max Planck Institute of Biochemistry, Department of Molecular Medicine, 82152, Martinsried, Germany

Received July 14, 2007; Revised November 27, 2007; Accepted November 30, 2007

## ABSTRACT

**Many *in vitro* and *in vivo* applications for transgenesis require co-expression of heterologous genes. The use of internal ribosome entry sites (IRESs) in dicistronic expression vectors enables the expression of two genes controlled by one promoter in target cells or whole organisms. Here we describe the expansion of IRES exploitation to generate multicistronic vectors capable of expressing multiple reporter genes, especially to improve the fluorescence yield of autofluorescent reporter gene products such as green fluorescent protein (GFP). We found that the increase in fluorescence output of GFP is proportional to the number of IRES-GFP repeats in the multicistronic vector. At least four genes can be expressed from a multicistronic vector by using tandem IRES elements, with no significant alteration of the expression level of the cap-dependent translated gene. Moreover, gene expression under the control of multiple IRES element has no effect on the posttranscriptional regulation through 3'-untranslated regions (3'UTR). Thus, endogenous gene expression and regulation, especially those controlled by weak promoters, can be analyzed with our IRES-dependent polycistronic reporter gene expression system.**

## INTRODUCTION

In 1988, Pelletier and Sonenberg (1) as well as Jang *et al.* (2) demonstrated that the internal ribosome entry site (IRES), originally found in eukaryotic viruses, initiate ribosome binding and translation, in a cap-independent manner. Since then, bicistronic vectors in which the first gene is translated in a cap-dependent manner and the second one in an IRES-dependent manner have been widely used for *in vitro* and *in vivo* applications in a variety of experimental settings, from cultured cells and

transgenic animals to gene therapy. For example, to link two genes transcribed from a single promoter, one being the gene of interest, and the other a selectable marker or a reporter gene (3–6). Furthermore, IRES-linked expression of fluorescent proteins has been used as reporter *in vivo* enabling to monitor the expression of target genes without deleting them (7–13).

However, there are limitations using IRES-linked reporter genes. Gene expression downstream of IRES is expressed at lower level as compared with the gene upstream of IRES (14–17), and fluorescence output using IRES-linked green fluorescent protein (GFP) expression is decreased or even absent compared with that found in GFP-fused proteins (18,19). Thus, the opportunity to expand IRES exploitation to generate multicistronic reporter vector capable of expressing multiple genes would open up new possibilities. For example, multiple IRES-enhanced GFP (EGFP) sequences linked in tandem could be used to improve the fluorescence yield. This can be valuable, especially for *in vivo* targeting of weakly expressed genes with IRES-controlled fluorescent marker.

Although some investigations have employed tagging proteins of interest using multicistronic constructs with multiple cloning sites (MCSs) flanking IRES elements (20–25), no quantitative and systematic studies of this approach have been performed yet. For example, (i) the quantity of the fluorescence output of a fluorescent protein expressed under the control of multiple tandem-IRES compared to single IRES is unknown; (ii) the effect of multiple tandem-IRES on the cap-dependent expressed gene is unclear and (iii) the influence of multiple tandem-IRES on the posttranscriptional regulation of gene expression is also unknown. In the present article we have addressed these questions.

## MATERIALS AND METHODS

### Vector construction

All vectors were constructed based on the vector backbone pIRES2-EGFP (BD Biosciences Clontech, Palo Alto, CA, USA; Catalog #6029-1) that contains the immediate early

\*To whom correspondence should be addressed. Tel: +49 89 51605437; Fax: +49 89 51605223; Email: bouabe@mvp.uni-muenchen.de

promoter of cytomegalovirus ( $P_{CMV}$  IE) followed by the IRES of the encephalomyocarditis virus (ECMV), an MCS, the enhanced GFP (EGFP) coding sequence and the SV40 polyadenylation signals (SV40pA) downstream of the EGFP.

The vectors were generated as follows. The luciferase coding sequence was amplified from pGL3-basic (Promega) using the forward primer CGCCTCGAGATGGAAGACGCCAAAACATAAAG containing XhoI restriction site and the backward primer CGCGAAT TCTTACACGGCGATCTTTCCGCCCTTC containing EcoRI restriction site. The product was purified, double digested with XhoI and EcoRI and inserted between the XhoI and EcoRI sites of pIRES2-EGFP. This construct was named pLuc-1x(IRES2EGFP)-SV40pA. The construct pLuc-2x(IRES2EGFP)-SV40pA was created as follows: the IRES2EGFP sequence was amplified from pIRES2EGFP plasmid using the forward primer CGC GCGGCCGCCCGCGGGCCCGGGATCCG containing NotI restriction site and the backward primer CGCGCG GCCGATATCTTACTTGTACAGCTCGTCCATGC CGAG containing NotI and EcoRV restriction sites. The insertion of the EcoRV restriction site allows cloning of further fragments downstream of IRES2EGFP, as described subsequently. The PCR-product was purified, digested with NotI and inserted into NotI site, between EGFP sequence and SV40pA, of pLuc-1x(IRES2EGFP)-SV40pA. To generate pLuc-3x(IRES2EGFP)-SV40pA, IRES2EGFP sequence was amplified from pIRES2EGFP using the forward primer CGCGATATCCCGCGGGCC CGGGATCCG and the backward primer CGCGATA TCTTACTTGTACAGCTCGTCCATGCCGAG, both containing EcoRV restriction sites. The PCR-product was purified, digested with EcoRV and inserted into the created EcoRV site, between the second EGFP sequence and SV40pA, of pLuc-2x(IRES2EGFP)-SV40pA.

To generate the constructs with the interleukin-10 (IL-10)-3' untranslated region (UTR) instead of SV40pA, we amplified a 832 base pair (bp) long IL-10-fragment located between the 6th and 838th bp behind the stop codon from mouse genomic DNA which enclose the 3'UTR with the forward primer GCGTCTAGAGTGCA GTGTGTATTGAGTCTGCTGG containing XbaI restriction site and the backward primer GCGCTTAAG CCGAGCTCAACCCCTTCCTGGAG containing AflIII restriction site. The SV40pA containing sequence was removed from the constructs pLuc-nx(IRES2EGFP)-SV40pA ( $n = 1, 2$  or  $3$ ) by digestion with XbaI and AflIII restriction sites. The XbaI/AflIII IL10-3'UTR fragment was then inserted the XbaI and AflIII of the pLuc-nx(IRES2EGFP) constructs. For the generation of the IL10-3'UTR containing constructs, to prevent Dam methylation of the Xba I sites, the constructs passed through a  $dam^- E. coli$  mutant (ATCC 47045).

### Cell culture and transfection

Prior to transfection, human embryonic kidney (HEK) 293 cell line were seeded in 6-well plates at  $5 \times 10^5$  cells per well in RPMI 1640 medium with L-glutamine (0.3 g/l),

10% fetal bovine serum (FCS), 10 mM HEPES and penicillin (100 U/ml)–streptomycin (100  $\mu$ g/ml) (Gibco/Invitrogen). The cells were allowed to attach overnight. A total of 0.9 pmol of each generated construct were transfected separately into HEK 293 using jetPEI<sup>TM</sup> transfection reagent (Qbiogene Molecular Biology). Selection with G418 (1 mg/ml) was started 48 h after transfection. Transfectants were obtained after 4–6 weeks and maintained under continued selection (0.4–1 mg/ml G418).

### Flow cytometry analysis

Cells were washed twice with phosphate buffered saline (PBS) (Gibco/Invitrogen) and harvested by using  $1 \times$  trypsin ethylene diamine tetra-acetic acid (trypsin–EDTA) (Gibco/Invitrogen). Equal cell numbers were analysed for EGFP fluorescence by a fluorescence-activated cell sorter (FACS, BD FACSCanto). Background fluorescence was determined using non-transfected cells. A total of 10 000 events were collected using list-mode format for each experiment. Cells were gated on EGFP-expressing population. Data were acquired and analyzed using FlowJo-8.4.6 Software (Tree Star Inc; San Carlos, CA, USA). The mean values (represent the peak integrals) of fluorescence for gated EGFP-expressing populations were determined.

### Western blot analysis

Equal numbers of cells ( $5-10 \times 10^6$ ) were washed twice in ice-cold PBS, resuspended in lysis-buffer (50 mM Tris–Hcl, pH 7.4; 150 mM NaCl; 1 mM EDTA; 2% chaps or 1% Triton X-100) and incubated on ice for 15–30 min. Cell debris was removed by centrifugation at 14 000g, 4°C for 10 min and the supernatants were transferred into new tubes. The protein concentration was determined by using BCA Protein Assay Kit (Pierce). Equal amounts of protein lysates were subjected to 12% SDS–PAGE under reducing conditions and transferred to nitrocellulose (NC) membrane by semi-dry electroblotting. The NC-membrane was blocked (PBS; 5% dry milk powder; 0.1% Tween-20) for 1–2 h at room temperature and incubated overnight at 4°C with anti-GFP or anti-actin antibodies (Santa Cruz Biotechnology). The antibodies were used in a final concentration of 0.4  $\mu$ g/ml. Bands were visualized with horseradish peroxidase-conjugated secondary antibody (Santa Cruz Biotechnology) and enhanced chemiluminescence substrate. All antibodies were diluted in PBS, 5% FCS and 0.1% Tween-20.

### Preparation of total RNA and cDNA synthesis

Total RNA was extracted from cells using the RNA MasterPure RNA Purification kit (Biozym). Contaminating genomic DNA was removed by treatment of the RNA preparations with Baseline-Zero DNase (Biozym). mRNA was stored at  $-80^\circ\text{C}$ . cDNA was generated from total RNA using the RevertAid<sup>TM</sup> H Minus First Strand cDNA Synthesis Kit (Fermentas), according to manufacturer's instructions. One microgram total RNA was used for each reaction and synthesis was primed by Oligo(dT)<sub>18</sub> primers. For negative controls, identical set-ups of reverse

transcriptase-polymerase chain reaction (RT-PCR) reactions except reverse transcriptase were performed. This negative control was performed to exclude products amplified from a potential genomic DNA contamination. One microliter of cDNA product was used for a subsequent PCR using 2×BioMix Red (Bioline) in a final volume of 30 µl. The expression level of the multicistronic vectors was detected after 30 amplification cycles using the following primer for luciferase: the forward primer, Luc-F: GGCGCTCCCCCTCTCTAAGGA and the backward primer, Luc-B: GGGGTGTTGGAGCAAGA TGG. As a loading control, the housekeeping gene, hypoxanthine-guanine-phosphoribosyltransferase (HPRT), was amplified using the primer set: HumHPRT-F: CAG CCCTGGCGTCGTGATTAGTGATG and HumHPRT-B: GGGAACTGATAGTCTATAGGCTCATAGTGC.

#### Preparation of genomic DNA and Southern blot analysis

Cells were seeded on 6-well plates and cultured until reaching complete confluence (~5 × 10<sup>6</sup> cells). Supernatants were then aspirated and 1.5 ml lysis buffer [100 mM Tris-HCl (pH 8.5); 0.2% SDS; 200 mM NaCl; 5 mM EDTA (pH 8.0); 100 µg/ml Proteinase K (freshly added)] was added to each well. After incubation at 37°C for at least 8 h, equal volume (1.5 ml) of isopropanol was added to each well and the plates were incubated on a shaker at 300 r.p.m. and RT for at least 3 h, until the DNA became visible. DNA were fished with pipette tip, transferred into 150 µl TE-buffer [10 mM Tris-HCl (pH 8); 1 mM EDTA (pH 8)] and incubated at 55°C overnight to allow the dissolution of DNA. The DNA was stored at room temperature (RT) until use. Forty microliter from each DNA sample was digested with 50 U EcoRI (Fermentas) overnight. The digested DNA samples were then fractionated on 0.7 agarose gel in 1×TBE (89 mM Tris-borate, 89 mM boric acid, 2 mM EDTA, pH 8). The gel was treated successively with depurination solution (0.2 M HCl), denaturing buffer (0.5 M NaOH; 1.5 M NaCl) and neutralization buffer (0.5 M Tris-HCl; 1.5 M NaCl; pH 7.2) for ~45 min, respectively. The gel was subsequently blotted overnight to a positive charged Nylon membrane (Schleicher and Schuell) using 20× SSC (3 M NaCl; 0.3 M tri-Sodium Citrate 2-hydrate; pH 7.2). Cross-linking of the DNA to the membrane was induced by irradiation with a UV lamp for 50 s. The membrane was incubated with hybridization solution [50% Formamide; 5× SSC; 10 mM Tris-HCl (pH 7.5); 1% SDS; 5× Denhardt's (2% Ficoll 400; 2% Polyvinylpyrrolidone K30; 2% BSA); 10% Dextran sulfate; 100 µg/ml salmon sperm DNA] for 1 h.  $\alpha^{32}\text{P}$ -dATP-labeled luciferase probe was then added to the hybridization solution and incubated overnight at 42°C. The luciferase probe was labeled using HexaLabel & trade DNA Labeling Kit (Fermentas) according to manufacturer's instructions. The blot was subsequently washed two times with wash-buffer 1 (2× SSC; 0.1% SDS) for 10 min at RT, then two times with wash buffer 2 (0.5× SSC; 0.4% SDS) for 20 min at 70°C. Radioactive signals were detected with Fuji Film Scanner FLA-3000 and analyzed with Aida Image Analyser v.4.00 software.

#### Luciferase assay

Equal numbers of cells were harvested, washed with PBS, resuspended in luciferase-lysis buffer (Luc-Buffer: 80 mM K<sub>2</sub>HPO<sub>4</sub>; 20 mM KH<sub>2</sub>PO<sub>4</sub>; 2 mM EDTA; 1% TritonX-100; pH adjusted to 7.8 with 1 M KOH; 1 mM DTT was added freshly before use) and incubated on ice for 15–30 min. Cell debris was removed by centrifugation at 14 000g, 4°C for 10 min and the supernatants were transferred into new tubes. The protein concentration was determined by using BCA Protein Assay Kit (Pierce). Prior to luciferase assay, bovine serum albumin (BSA) was added to a final concentration of 0.5%. The cell lysates were stored at –20°C. Equal amounts of protein lysates (10–60 µg) were transferred into a white microtiter plate (volumes were adjusted to 50–100 µl with Luc-Buffer/0.5% BSA). One volume of Luciferase Assay Reagent (20 mM Tricine; 1.07 mM (MgCO<sub>3</sub>)<sub>4</sub>Mg(OH)<sub>2</sub>·5H<sub>2</sub>O; 0.1 mM EDTA; pH adjusted to 7.8 with 1 M HCl; 33.3 mM DTT; 270 µM Li<sub>3</sub>-Coenzyme A; 470 µM D-luciferin; 530 µM Mg-ATP) was added to the cell lysate and chemiluminescence was measured immediately with a MicroLumatPlus LB96V microplate luminometer (EG&G Berthold). All experiments were performed independently at least three times.

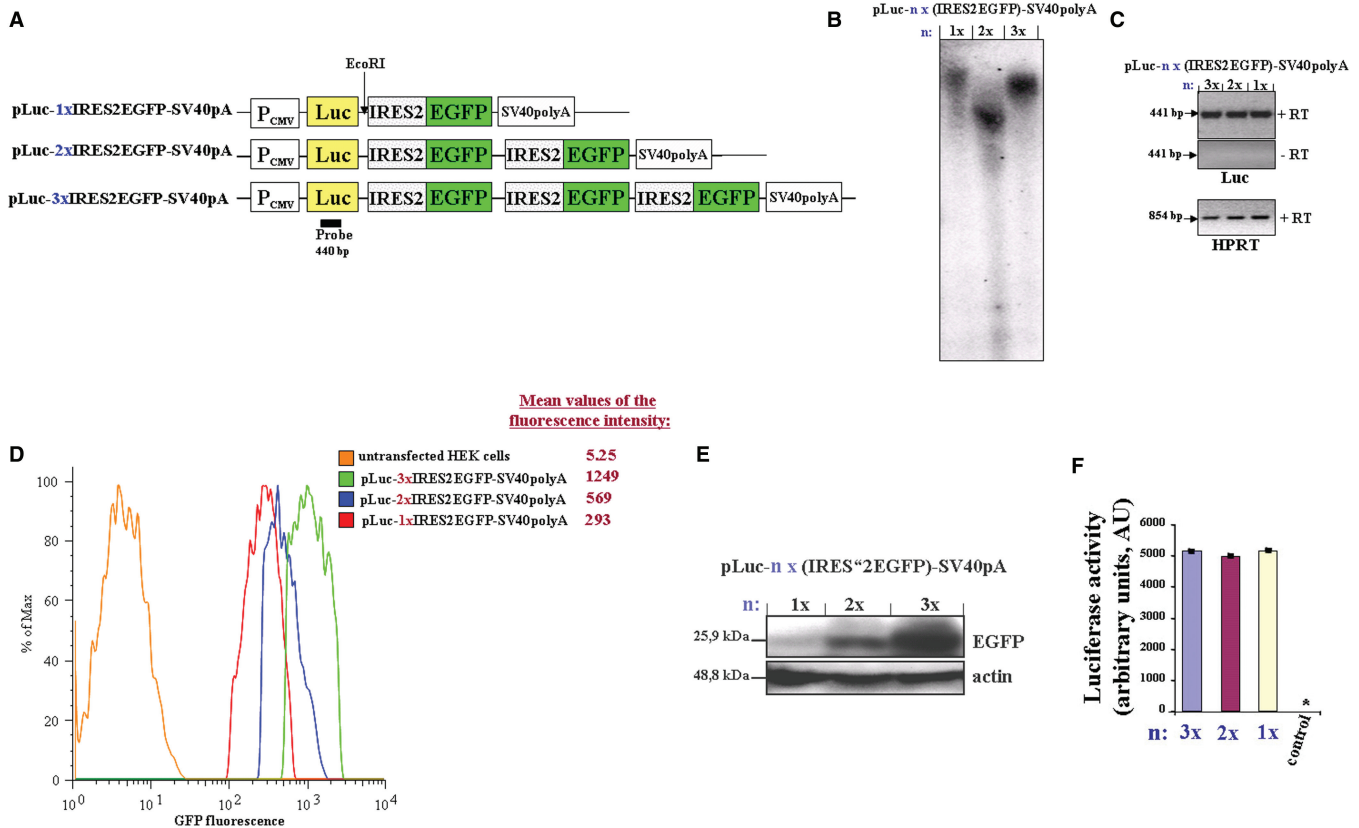
## RESULTS

#### Improvement of the fluorescence yield by expressing EGFP under the control of multiple tandem-IRES

To test whether an increase in fluorescence output of EGFP can be achieved by expressing EGFP from a multicistronic vector using the IRES of the ECMV in tandem, three constructs, named pLuc-nx(IRES2-EGFP)-SV40-pA (*n*=1, 2 or 3), were generated which contain one, two or three IRES-EGFP repeats (Figure 1A), respectively and stably expressed in HEK 293 cells.

To determine the copy number of inserted constructs, genomic DNA was isolated from the HEK 293 cell lines, digested with EcoRI and subjected to Southern blot analysis using a luciferase probe. The pLuc-nx(IRES2-EGFP)-SV40-pA constructs contain only one EcoRI restriction site, directly downstream of the luciferase gene (Figure 1A). The Southern blot analysis showed a single insertion of each construct into the genome of the cell lines (Figure 1B). Furthermore, the analysis of mRNA expression by RT-PCR showed no significant difference in the expression level of the three constructs (Figure 1C), as well no loss in the integrity of the mRNA, at least for the tetracistronic vector (data not shown). Thus, the fluorescence output of EGFP will be dependent on the number of IRES-EGFP repeats in each construct and on their translation efficiency and not on different genomic integration efficiency.

Flow cytometry analysis of the established three cell lines revealed a linear amplification of spectral yield of EGFP (Figure 1D). The mean fluorescence intensity for EGFP (Figure 1D and data not shown) showed that the increase in fluorescence output is proportional to the number of IRES-EGFP repeats. Western blot analysis of



**Figure 1.** (A) Schematic representation of the first three constructs used in the studies. The expression vectors were constructed as described in 'Materials and methods' section. The EcoRI restriction site and luciferase probe used for Southern blot analysis are shown. The constructs pLuc-nx(IRES2-EGFP)-SV40-pA ( $n = 1, 2$  or  $3$ ) were stably transfected in HEK293 cells. (B) Southern blot analysis shows a single insertion of each construct into the genome of the generated three cell lines. (C) The expression level of the multicistronic RNA from the different constructs was analyzed by RT-PCR for luciferase (Luc). (+RT) and (-RT) refer to PCR product from first-strand cDNA synthesis reaction performed from Baseline-Zero DNase-treated total RNA with (+RT) or without (-RT, negative control) reverse transcriptase, respectively. As a loading control, RT-PCR of the housekeeping gene, HPRT, was used. (D) The fluorescence emission of EGFP were analyzed by flow cytometry (using BD FACSCanto). Shown are histograms gated on EGFP-positive cells. The mean values (peak integrals) of fluorescence for the gated EGFP-expressing cells are indicated. (E) Equal amounts of whole cell lysates were separated by SDS-PAGE and the expression level of EGFP were analyzed by western blot using EGFP-antibody. Actin immunostaining is shown as the loading control. (F) Equal amounts of whole cell lysates (10–60  $\mu$ g) were analyzed for luciferase activity. Control: non transfected HEK293 cells, \*, value <50 AU. Experiments were repeated at least three times and similar results were obtained. P<sub>CMV</sub>, cytomegalovirus promoter; Luc, luciferase gene; SV40polyA, Simian virus polyadenylation site.

whole cell lysates from those cell lines confirmed the increase of EGFP-expression (Figure 1E).

### Multiple tandem-IRES do not affect the upstream cap-dependent gene expression

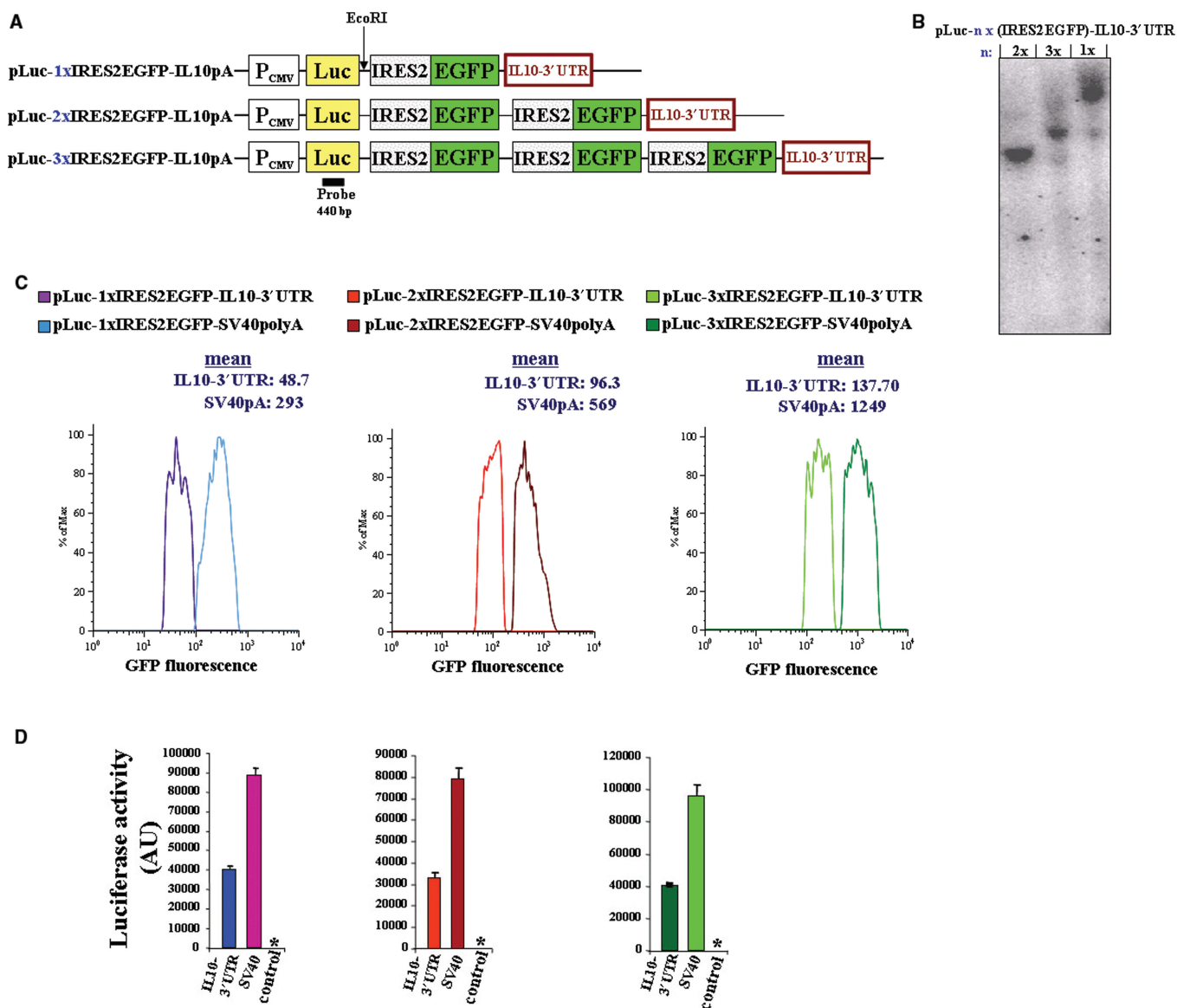
A gene encoding luciferase was inserted right after P<sub>CMV</sub> and upstream of the first IRES into the three constructs (pLuc-nx(IRES2-EGFP)-SV40-pA; Figure 1A) in order to quantitatively analyze cap-dependent gene expression. If the IRES-dependent translations would interfere with the cap-dependent translation, the luciferase activity should be reciprocated proportional to the number of inserted IRES-EGFP repeats. Equal amounts of whole cell lysates from the three cell lines were collected and analysed for luciferase activity. Interestingly, the luciferase reporter activities were similar for the three different constructs (Figure 1F) indicating that multiple tandem-IRES do not interfere with cap-dependent gene expression. Since the cap-dependent luciferase activity is not affected by the

number of IRESs, it can be used as an internal normalization control to eliminate experimental variability, including variations in transfection efficiency.

### Multiple tandem-IRES do not influence the posttranscriptional regulation through sequences in the 3'UTRs

The 'adenylate/uridylylate (AU)-rich elements' (AREs) are well-known RNA sequences in 3'UTRs, which regulate the stability of the mRNA and hence the expression level of the corresponding protein (26). We therefore asked whether the presence of several exogenous IRES sequences in the 3' region of mRNA would interfere with the 3'UTR function.

To test this, we replaced the RNA-stabilizing SV40polyA by the IL-10-3'UTR of the pLuc-nx(IRES2-EGFP)-SV40-pA constructs (Figure 2A). IL-10-3'UTR is known to accelerate the degradation of IL-10-mRNA (27). We used a 832 bp long mouse IL-10-fragment



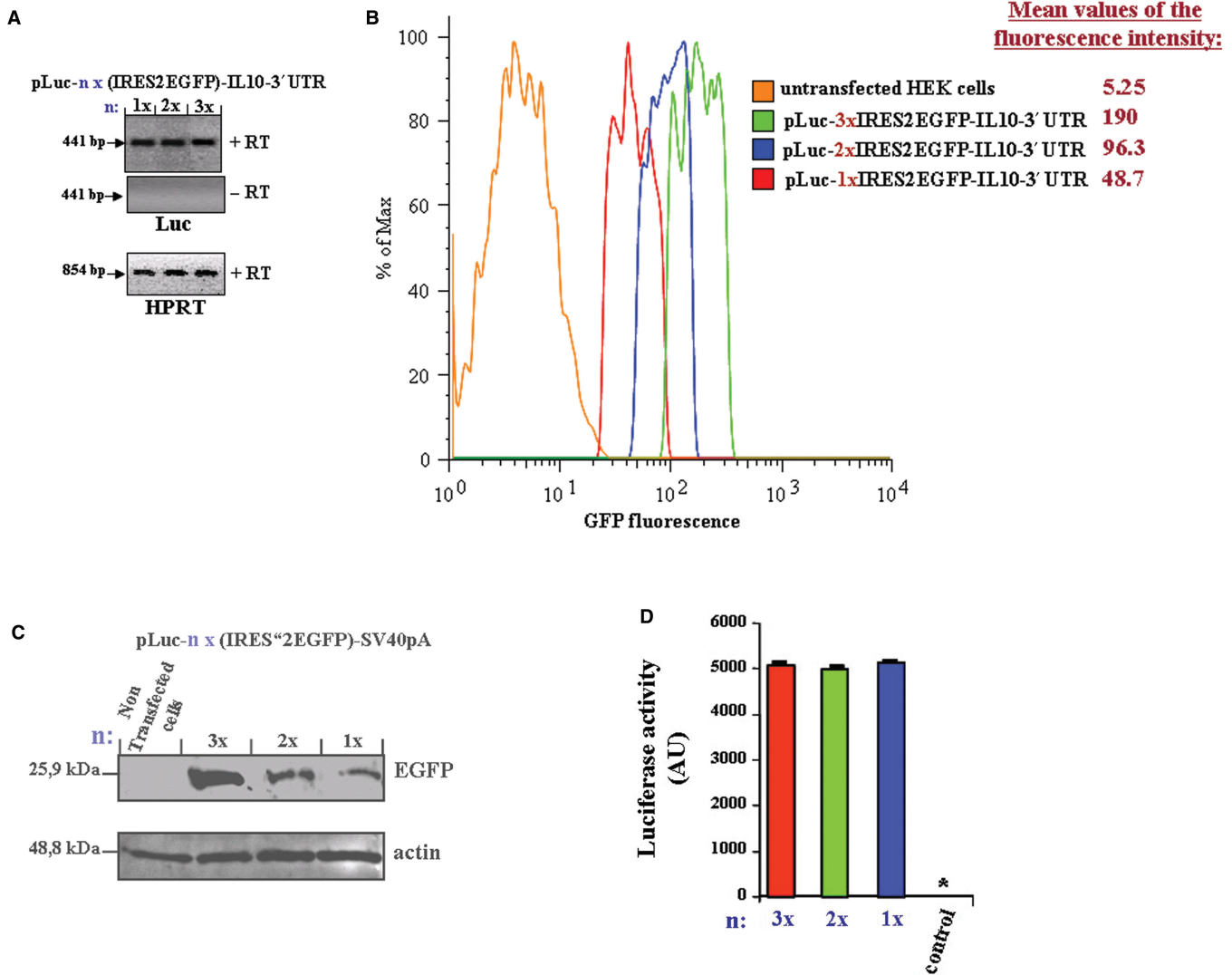
**Figure 2.** (A) Schematic representation of the pLuc-nx(IRES2-EGFP)-IL10-3'UTR constructs ( $n = 1, 2$  or  $3$ ). The expression vectors were constructed as described in 'Materials and methods' section. The EcoRI restriction site and luciferase probe used for Southern blot analysis are shown. The constructs were stably transfected in HEK293 cells. (B) Southern blot analysis shows a single insertion of each construct into the genome of the generated three cell lines. (C) Comparison of the fluorescence emission of EGFP between SV40polyA and IL10-3'UTR-containing constructs were determined by flow cytometry. Shown are histograms gated on EGFP-positive cells. The mean values (peak integrals) of fluorescence for the gated EGFP-expressing cells are indicated. (D) Equal amounts of whole cell lysates from all six cell lines expressing the indicated construct were analyzed for luciferase activity. Control, nontransfected HEK293 cells; \*, value  $< 15$  AU. P<sub>CMV</sub>, cytomegalovirus promoter; SV40polyA, Simian virus polyadenylation site.

spanning the 6th and 838th bp downstream of the stop codon, which include the 3'UTR sequence.

To determine the copy number of inserted constructs, genomic DNA was isolated from the HEK 293 cell lines, digested with EcoRI and subjected to Southern blot analysis using a luciferase probe. Similar to the pLuc-nx(IRES2-EGFP)-SV40-pA expressing cell lines, the Southern blot analysis showed a single insertion of each construct into the genome of the pLuc-nx(IRES2-EGFP)-IL10-3'UTR expressing cell lines (Figure 2B). Thus, the expression level of the different constructs will be

dependent on posttranscriptional regulation through the sequences in the 3'-untranslated region.

Flow cytometry analysis of the cell lines stably expressing multicistronic vectors containing IL-10-3'UTR showed decreased fluorescence intensity compared to SV40polyA containing constructs (Figure 2C). The insertion of the luciferase gene upstream of the IRES allowed a second readout system for gene expression monitoring luciferase reporter activities. Exchange of SV-40polyA for an IL-10-3'UTR, results in the loss of luciferase activities by more than 50% (Figure 2D). This is in good agreement



**Figure 3.** Stably transfected HEK293 cells with the constructs pLuc-nx(IRES2-EGFP)-IL10-3'UTR, containing one ( $n = 1$ ), two ( $n = 2$ ) or three ( $n = 3$ ) IRES-EGFP repeats were established and (A) The expression level of the multicistronic RNA from the different constructs was analyzed by RT-PCR. (+RT) and (-RT) refer to PCR product from first-strand cDNA synthesis reaction performed from Baseline-Zero DNase-treated total RNA with (+RT) or without (-RT, negative control) reverse transcriptase (RT), respectively. As a loading control, RT-PCR of the housekeeping gene, HPRT, was used. (B) The fluorescence emission of EGFP from the three cell lines were analyzed by flow cytometry. Shown are histograms gated on EGFP-positive cells. The mean values (peak integrals) of fluorescence for the gated EGFP-expressing cells are indicated. (C) Equal amounts of whole cell lysates were separated by SDS-PAGE and the expression level of EGFP were analyzed by western blot using EGFP-antibody. Actin immunostaining is shown as the loading control. (D) Equal amounts of whole cell lysates from all three cell lines were analyzed for luciferase activity. Control, nontransfected HEK293 cells; \*, value <15 AU.

with results by Powell *et al.* (27). Thus, multiple tandem-IRES do not interfere with the posttranscriptional regulation of gene expression through sequences in the 3'UTR.

#### The proportional increase in fluorescence intensity of EGFP is independent of the posttranscriptional regulation by 3' UTR

With the pLuc-nx(IRES2-EGFP)-SV40-polyA constructs the increase in fluorescence output was proportional to the number of IRES-EGFP repeats (Figure 1D). We wondered whether similar result can be reached with the IL10-3'UTR containing constructs, which show an altered

expression regulation compared to SV40polyA containing constructs (Figure 2).

The analysis of mRNA expression by RT-PCR showed no significant difference in the expression level of the three different constructs (Figure 3A). Furthermore, the Southern blot analysis showed a single insertion of each construct into the genome of the pLuc-nx(IRES2-EGFP)-IL10-3'UTR expressing cell lines (Figure 2B). Thus, the fluorescence output of EGFP will be dependent on the number of IRES-EGFP repeats and on their translation efficiency and not on different genomic integration efficiency.

Flow cytometry analysis of the established HEK293 cell lines, expressing multicistronic vectors containing

IL10-3'UTR and one, two or three IRES-EGFP repeats, respectively, showed a proportional increase in fluorescence output, dependent on the number of IRES-EGFP repeats (Figure 3B). Western blot analysis of whole cell lysates from those three cell lines confirmed the increase of EGFP-Expression (Figure 3C).

Furthermore, the luciferase reporter activities that are cap-dependent, were similar between the cell lysates from the three cell lines (Figure 3D).

Thus, independent of the posttranscriptional regulation through sequences in the 3'-UTR, the increase in fluorescence output of EGFP is proportional to the number of IRES-EGFP repeats, without altering the cap-dependent gene expression.

## DISCUSSION

Viral IRES elements can be fused with target genes resulting in the coexpression of multiple proteins from the same mRNA (3–6,28). However, cistrons downstream of IRES are often much less efficiently translated than those directly following the promoter (14–16 and Bouabe, H., unpublished data), and subsequently IRES-directed gene expression using nonenzymatic fluorescent reporters, such as EGFP, result in reduced or even absent fluorescence emission (18,19,29,30). If gene expression from weak promoter should be monitored, the IRES-linked fluorescent reporter gene is usually not ideal.

Here we report a technology that allows improvement of reporter gene activity using multiple IRES-EGFP sequences linked in tandem. Di-, tri- and tetra-cistronic vectors were constructed to express simultaneously luciferase (Luc) gene in a cap-dependent and one, two or three EGFP genes in an IRES-dependent manner, respectively. We used an EMCV-IRES (BD Biosciences Clontech). Flow cytometry and western blot analysis demonstrated linear amplification of fluorescence signal and expression level of EGFP, respectively, which were proportional to the number of IRES-EGFP repeats. Thus, to compensate the IRES-dependent decreased expression, several copies of a gene of interest can be expressed from a multicistronic vector by using tandem IRES elements.

Our results also indicate that all three cistrons following the IRESs are equally translated. Thus, there is no significant influence of each of the IRES-mediated translation on the translation efficiency mediated by the other IRES. Similar results are reported by other observations using di- and tri-cistronic vectors (31,32). Thus the translations of all three EGFP cistrons downstream of IRESs may occur independent of each other.

Translation of capped cellular mRNAs is initiated by recruitment of ribosomes to the 5' end of mRNAs via eukaryotic translation initiation factors (eIF)4E, eIF4A and eIF4G. Translation initiation on most IRES-containing mRNAs, such as EMCV mRNA, requires the same canonical eIFs that are required for translation of capped mRNAs, except for eIF4E (33). It is suggested that IRES containing EMCV mRNA competes with capped cellular mRNA for the recruitment of eIF4G and subsequently for translation (34). This raises the

question, whether multiple tandem-IRES in multicistronic RNA affect the cap-dependent gene expression.

Recently, a generated knock-in mouse model, where an IRES-luciferase-IRES-EGFP cassette was inserted into the UTR of the endogenous transcription factor forkhead box P3 (Foxp3) locus, resulting in attenuated expression of the cap-dependent translated endogenous Foxp3 gene (35). The authors assumed that four AREs found dispersed in the luciferase complementary DNA caused the destabilization of the Foxp3 mRNA and subsequently reduced Foxp3 expression.

However, under our experimental conditions, whether one, two or three IRES-EGFP cassettes are inserted downstream of the Luc-gene, the cap-dependent luciferase expression remains similar. Thus, multiple tandem-IRES do not affect the cap-dependent gene expression. In line with our results, Hennecke *et al.* (36) reported that cap-dependent translation from bicistronic mRNAs remains comparable to monocistronic expression. Nevertheless, it remains to be elucidated, whether this is also valid for other cell types, and in particular for cells under stress conditions, which are known to modulate the availability of eIF4E for cap-dependent translation (34,37)

Posttranscriptional events are central to the regulation of eukaryotic gene expression [reviewed in (38,39)]. They act on the modulation of mRNA stability and/or translation efficiency via the binding of heterogeneous nuclear ribonucleoproteins (hnRNPs) to specific RNA sequences, which are often present in the 3'-UTR (40–42). The AREs are the best-known RNA sequences of the 3'UTR (43). They can bind diverse proteins that either stabilize the mRNA, such as HuR, or accelerate their degradation, such as AUF1 protein and Tristetraprolin (26,44,45), thereby enhancing or decreasing the corresponding translation. Thus, considering on one hand several exogenous IRES sequences present in the 3' region of the mRNA, which recruit and bind diverse proteins and ribosome complexes, and on the other hand that cellular 3'UTR and IRES are modulated, in part, by common RNA-binding proteins, like polypyrimidine tract-binding protein (PTB) (5,46,47), it seems very likely that multiple tandem-IRES in an mRNA may influence its posttranscriptional regulation through 3'UTR.

We therefore compared the effect of multiple IRESs on the 3'UTR-mediated posttranscriptional regulation by inserting an IL-10-3'UTR as RNA-destabilizing fragment, instead of the SV-40polyA in the pLuc-nx(IRES2-EGFP) constructs. Whether one, two or three IRES-EGFP cassettes are inserted upstream of the 3'UTR, the IL-10-3'UTR retains its ability to decrease the expression level of EGFP and luciferase. Thus, multiple tandem-IRES do not affect the posttranscriptional regulation of gene expression through sequences in the 3'UTRs.

Several studies have established that the IRES element of EMCV is functional in a variety of cell types and in whole animals. Using gene targeting strategies several reporter mice have been generated, where the expression of a gene of interest is linked via IRES element with EGFP. These mice enable us to determine which cells express the target gene *in vivo* (7–13). However, there are several limitations using IRES-linked EGFP knock-in mice.

Only some of them are mentioned subsequently. First, IRES-dependent gene expression ranged between 20% and 50% that of the cap-dependent gene (18 and Bouabe, H., unpublished data). Second, in contrast to transgenic animals with several copies of a transgene supplied by retroviruses or microinjection, there are only one or two copies of the IRES-EGFP reporter gene in the genome of knock-in mice generated by gene targeting, thus the IRES-dependent decreased expression cannot be compensated for by high copy number of the transgene. Third, the expression level of the EGFP reporter is dependent on the cell- and tissue-types, as well as on the promoter capacity of the target gene. For these reasons and in view of EGFP's limit of detection that is approximately  $10^5$  molecules/cell (48), care should be taken regarding the decreased capacity of IRES-dependent EGFP reporter gene expression. It is very likely that IRES-linked expression of EGFP would not mark all cells expressing the target gene, especially that expressed under weak promoter. Subsequently, the reporter mice would provide incomplete gene expression profile.

In this study, we provide a reporter system that can help to overcome these limitations, at least to some extent, by using multiple IRES-EGFP sequences linked in tandem. Multiple tandem-IRES can enhance the reporter gene activity with no significant effect on the expression regulation of the target gene upstream of IRES. Thus, they perform essential prerequisites allowing their use for *in vivo* studies.

To evaluate the multiple tandem-IRES-based reporter system *in vivo*, we used the gene targeting strategy to integrate multicistronic vectors in a specific genomic site of embryonic stem (ES) cells. The reporter mice that we are generating from these modified ES-cell lines, will allow study of the expression behavior of IRES-mediated polycistronic reporter system *in vivo*.

Finally, in biochemical and cellular studies, a defined expression level of the genes of interest is required. Overexpression can lead to aggregation or to modified function and/or subcellular localization of the protein of interest. IRES-dependent polycistronic gene expression opens up the possibility, to titrate the protein expression level by expressing the gene of interest under the optimal IRES-tandem repeats.

## ACKNOWLEDGEMENTS

We are grateful to Dr Cemalettin Bekpen for critical reading of the manuscript. Hicham Bouabe was supported by the DFG Graduate College "Infection and Immunity". Funding to pay the Open Access publication charges for this article was provided by Max von Pettenkofer Institut, Germany.

*Conflict of interest statement.* None declared.

## REFERENCES

1. Pelletier, J. and Sonenberg, N. (1988) Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*, **334**, 320–325.

2. Jang, S.K., Krausslich, H.G., Nicklin, M.J.H., Duke, G.M., Palmenberg, A.C. and Wimmer, E. (1988) A Segment of the 5' nontranslated region of encephalomyocarditis virus RNA directs internal entry of ribosomes during *in vitro* translation. *J. Virol.*, **62**, 2636–2643.
3. Mountford, P.S. and Smith, A.G. (1995) Internal ribosome entry sites and dicistronic RNAs in mammalian transgenesis. *Trends Genet.*, **5**, 179–184.
4. Gallardo, H.F., Tan, C. and Sadelain, M. (1997) The internal ribosomal entry site of the encephalomyocarditis virus enables reliable coexpression of two transgenes in human primary T lymphocytes. *Gene Ther.*, **10**, 1115–1119.
5. Martinez-Salas, E. (1999) Internal ribosome entry site biology and its use in expression vectors. *Curr. Opin. Biotechnol.*, **5**, 458–464.
6. Attal, J., Theron, M.C. and Houdebine, L.M. (1999) The optimal use of IRES (internal ribosome entry site) in expression vectors. *Genet. Anal.*, **15**, 161–165.
7. Mohrs, M., Shinkai, K., Mohrs, K. and Locksley, R.M. (2001) Analysis of type 2 immunity *in vivo* with a Bicistronic IL-4 Reporter. *Immunity*, **15**, 303–311.
8. Lorsbach, R.B., Moore, J., Ang, S.O., Sun, W., Lenny, N. and Downing, J.R. (2004) Role of RUNX1 in adult hematopoiesis: analysis of RUNX1-IRES-GFP knock-in mice reveals differential lineage expression. *Blood*, **7**, 2522–2529.
9. Im, W., Kim, H., Yun, D., Seo, S.Y., Park, S.H., Locksley, R.M. and Hong, S. (2005) Cytokine reporter mouse system for screening Novel IL12/23 p40-inducing compounds. *Mol. Cells*, **20**, 288–296.
10. Mohrs, K., Wakil, A.E., Killeen, N., Locksley, R.M. and Mohrs, M. (2005) A two-step process for cytokine production revealed by IL-4 dual-reporter mice. *Immunity*, **23**, 419–429.
11. Wan, Y.Y. and Flavell, R.A. (2005) Identifying Foxp3-expressing suppressor T cells with a bicistronic reporter. *PNAS*, **102**, 5126–5131.
12. Reinhardt, R.L., Hong, S., Kang, S.J., Wang, Z. and Locksley, R.M. (2006) Visualization of IL-12/23p40 *in vivo* reveals immunostimulatory dendritic cell migrants that promote Th1 differentiation. *J. Immunol.*, **177**, 1618–1627.
13. Kamanaka, M., Kim, S.T., Wan, Y.Y., Sutterwala, F.S., Lara-Tejero, M., Galan, J.E., Harhaj, E. and Flavell, R.A. (2006) Expression of interleukin-10 in intestinal lymphocytes detected by an interleukin-10 reporter knockin tiger mouse. *Immunity*, **25**, 941–952.
14. Kaufman, R.J., Davies, M.V., Wasley, L.C. and Michnick, D. (1991) Improved vectors for stable expression of foreign genes in mammalian cells by use of the untranslated leader sequence from EMC virus. *Nucleic Acids Res.*, **16**, 4485–4490.
15. Houdebine, L.M. and Attal, J. (1999) Internal ribosome entry sites (IRESs): reality and use. *Transgenic Res.*, **3**, 157–177.
16. Dirks, W., Wirth, M. and Hauser, H. (1993) Dicistronic transcription units for gene expression in mammalian cells. *Gene*, **2**, 247–249.
17. Zhou, Y., Aran, J., Gottesman, M.M. and Pastan, I. (1998) Co-expression of human adenosine deaminase and multidrug resistance using a bicistronic retroviral vector. *Hum. Gene Ther.*, **9**, 287–293.
18. Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E. and Hayakawa, T. (2000) IRES-dependent second gene expression is significantly lower than cap-dependent first gene expression in a bicistronic vector. *Mol. Ther.*, **1**, 376–382.
19. Yu, X., Zhan, X., D'Costa, J., Tanavde, V.M., Ye, Z., Peng, T., Malehorn, M.T., Yang, X., Civin, C.I. *et al.* (2003) Lentiviral vectors with two independent internal promoters transfer high-level expression of multiple transgenes to human hematopoietic stem-progenitor cells. *Mol. Ther.*, **7**, 827–838.
20. Fussenegger, M., Mazur, X. and Bailey, J.E. (1998) pTRIDENT, a novel vector family for tricistronic gene expression in mammalian cells. *Biotechnol. Bioeng.*, **1**, 1–10.
21. Fussenegger, M., Moser, S. and Bailey, J.E. (1998) pQuattro vectors allow one-step multigene metabolic engineering and auto-selection of quattrocistronic artificial mammalian operons. *Cytotechnology*, **7**, 229–235.
22. Fussenegger, M., Moser, S., Mazur, X. and Bailey, J.E. (1997) Autoregulated multicistronic expression vectors provide one-step



- cloning of regulated product gene expression in mammalian cells. *Biotechnol. Prog.*, **6**, 733–740.
23. Moser, S., Schlatter, S., Fux, C., Rimann, M., Bailey, J.E. and Fussenegger, M. (2000) An update of pTRIDENT multicistronic expression vectors: pTRIDENTs containing novel streptogramin-responsive promoters. *Biotechnol. Prog.*, **5**, 724–235.
  24. Azzouz, M., Martin-Rendon, E., Barber, R.D., Kyriacos, A., Mitrophanous, K.A., Carter, E.E., Rohll, J.B., Kingsman, S.M., Kingsman, A.J. *et al.* (2002) Multicistronic lentiviral vector-mediated striatal gene transfer of aromatic L-amino acid decarboxylase, tyrosine hydroxylase, and GTP cyclohydrolase I induces sustained transgene expression, dopamine production, and functional improvement in a rat model of Parkinson's disease. *J. Neurosci.*, **23**, 10302–10312.
  25. Li, T. and Zhang, J. (2004) Stable expression of three genes from a tricistronic retroviral vector containing a picornavirus and 9-nt cellular internal ribosome entry site elements. *J. Virol. Methods*, **2**, 137–144.
  26. Brennan, C.M. and Steitz, J.A. (2001) HuR and mRNA stability. *Cell Mol. Life Sci.*, **2**, 266–277.
  27. Powell, M.J., Thompson, S.A., Tone, Y., Waldmann, H. and Tone, M. (2000) Posttranscriptional regulation of IL-10 gene expression through sequences in the 3'-untranslated region. *J. Immunol.*, **1**, 292–296.
  28. Fussenegger, M., Bailey, J.E., Hauser, H. and Mueller, P.P. (1999) Genetic optimization of recombinant glycoprotein production by mammalian cells. *Trends Biotechnol.*, **1**, 35–42.
  29. Ward, C.M. and Stern, P.L. (2002) The human cytomegalovirus immediate-early promoter is transcriptionally active in undifferentiated mouse embryonic stem cells. *Stem Cells*, **5**, 472–475.
  30. Chung, S., Andersson, T., Sonntag, K.C., Bjorklund, L., Isacson, O. and Kim, K.S. (2002) Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines. *Stem Cells*, **2**, 139–145.
  31. Zhu, J., Musco, M.L. and Grace, M.J. (1999) Three-color flow cytometry analysis of tricistronic expression of eBFP, eGFP, and eYFP using EMCV-IRES linkages. *Cytometry*, **1**, 51–59.
  32. Li, J., Menzel, C., Meier, D., Zhang, C., Dubel, S. and Jostock, T. (2007) A comparative study of different vector designs for the mammalian expression of recombinant IgG antibodies. *J. Immunol. Methods*, **318**, 113–124.
  33. Pestova, T.V., Hellen, C.U. and Shatsky, I.N. (1996) Canonical eukaryotic initiation factors determine initiation of translation by internal ribosomal entry. *Mol. Cell Biol.*, **12**, 6859–6869.
  34. Svitkin, Y.V., Herdy, B., Costa-Mattioli, M., Gingras, A.C., Raught, B. and Sonenberg, N. (2005) Eukaryotic translation initiation factor 4E availability controls the switch between cap-dependent and internal ribosomal entry site-mediated translation. *Mol. Cell Biol.*, **23**, 10556–10565.
  35. Wan, Y.Y. and Flavell, R.A. (2007) Regulatory T-cell functions are subverted and converted owing to attenuated Foxp3 expression. *Nature*, **7129**, 766–770.
  36. Hennecke, M., Kwissa, M., Metzger, K., Oumard, A., Kröger, A., Schirmbeck, R., Reimann, J. and Hauser, H. (2001) Composition and arrangement of genes define the strength of IRES-driven translation in bicistronic mRNAs. *Nucleic Acids Res.*, **16**, 3327–3334.
  37. Richter, J.D. and Sonenberg, N. (2005) Regulation of cap-dependent translation by eIF4E inhibitory proteins. *Nature*, **7025**, 477–480.
  38. Hazelrigg, T. (1998) The destinies and destinations of RNAs. *Cell*, **4**, 451–460.
  39. Moore, M.J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science*, **5740**, 1514–1518.
  40. Blaxall, B.C., Pende, A., Wu, S.C. and Port, J.D. (2002) Correlation between intrinsic mRNA stability and the affinity of AUF1 (hnRNP D) and HuR for A+U-rich mRNAs. *Mol. Cell Biochem.*, **232**, 1–11.
  41. Bevilacqua, A., Ceriani, M.C., Capaccioli, S. and Nicolin, A. (2003) Post-transcriptional regulation of gene expression by degradation of messenger RNAs. *J. Cell. Physiol.*, **3**, 356–372.
  42. Chen, C.Y., Xu, N., Zhu, W. and Shyu, A.B. (2004) Functional dissection of hnRNP D suggests that nuclear import is required before hnRNP D can modulate mRNA turnover in the cytoplasm. *RNA*, **4**, 669–680.
  43. Chen, C.Y. and Shyu, A.B. (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem. Sci.*, **11**, 465–470.
  44. Malter, J.S. (1989) Identification of an AUUUA-specific messenger RNA binding protein. *Science*, **4930**, 664–666.
  45. Wilusz, C.J., Wormington, M. and Peltz, S.W. (2001) The cap-to-tail guide to mRNA turnover. *Nat. Rev. Mol. Cell Biol.*, **2**, 237–246.
  46. Stoneley, M. and Willis, A.E. (2004) Cellular internal ribosome entry segments: structures, trans-acting factors and regulation of gene expression. *Oncogene*, **18**, 3200–3207.
  47. Martinez-Salas, E., Ramos, R., Lafuente, E. and Lopez de Quinto, S. (2001) Functional interactions in internal translation initiation directed by viral and cellular IRES elements. *J. Gen. Virol.*, **5**, 973–984.
  48. Tsien, R.Y. (1998) The green fluorescent protein. *Annu. Rev. Biochem.*, **67**, 509–544.