STEM CELLS®

EMBRYONIC STEM CELLS/INDUCED PLURIPOTENT STEM CELLS

Applying a "Double-Feature" Promoter to Identify Cardiomyocytes Differentiated from Human Embryonic Stem Cells Following Transposon-Based Gene Delivery

TAMÁS I. ORBÁN,[a] ÁGOTA APÁTI,[a] ANDREA NÉMETH,[a] NÓRA VARGA,[a] VIRÁG KRIZSIK,[a] ANITA SCHAMBERGER,[a] KORNÉLIA SZEBÉNYI,[a] ZSUZSA ERDEI,[a] GYÖRGY VÁRADY,[a] ÉVA KARÁSZI,[a] LÁSZLÓ HOMOLYA,[a] KATALIN NÉMET,[a] ELEN GÓCZA,[b] CSABA MISKEY,[c] LAJOS MÁTÉS,[c] ZOLTÁN IVICS,[c] ZSUZSANNA IZSVÁK,[c] BALÁZS SARKADI[a]

[a]Membrane Research Group of the Hungarian Academy of Sciences, Semmelweis University and National Blood Center, Budapest, Hungary; [b]Genetic Modification Program Group, Agricultural Biotechnology Center, Gödöllő, Hungary; [c]Mobile DNA Group, Max-Delbrück Center for Molecular Medicine, Berlin, Germany

Key Words. Sleeping Beauty transposon • Human embryonic stem cells • CAG promoter • ''Double-feature'' promoter • Cardiomyocytes • Lentiviral gene delivery

ABSTRACT

Human embryonic stem (HuES) cells represent a new potential tool for cell-therapy and gene-therapy applications. However, these approaches require the development of efficient, stable gene delivery, and proper progenitor cell and tissue separation methods. In HuES cell lines, we have generated stable, enhanced green fluorescent protein (EGFP)-expressing clones using a transposon-based (*Sleeping Beauty*) system. This method yielded high percentage of transgene integration and expression. Similarly to a lentiviral expression system, both the undifferentiated state and the differentiation pattern of the HuES cells were preserved. By using the CAG promoter, in contrast to several other constitutive promoter sequences (such as CMV, elongation factor 1α, or phosphoglycerate kinase), an exceptionally high EGFP expression was observed in differentiated cardiomyocytes. This phenomenon was independent of the transgene sequence, methods of gene delivery, copy number, and the integration sites. This "double-feature" promoter behavior, that is providing a selectable marker for transgene expressing undifferentiated stem cells, and also specifically labeling differentiated cardiomyocytes, was assessed by transcriptional profiling. We found a positive correlation between CAG promoter-driven EGFP transcription and expression of cardiomyocyte-specific genes. Our experiments indicate an efficient applicability of transposon-based gene delivery into HuES cells and provide a novel approach to identify differentiated tissues by exploiting a nontypical behavior of a constitutively active promoter, thereby avoiding invasive drug selection methods. STEM CELLS *2009;27:1077–1087*

Disclosure of potential conflicts of interest is found at the end of this article.

INTRODUCTION

The application of human embryonic stem (HuES) cells provides new hopes in the clinical treatment of a number of diseases and, at the same time, these stem cells are excellent models for studying tissue development and physiological cell differentiation. The so called "regenerative medicine" makes use of cells that can grow and differentiate to replace a damaged tissue. Also, efficient and stable gene delivery into stem cells should form the basis of successful gene therapy applications. Currently, the most widely applied methods for gene delivery into various stem/progenitor cells are based on the use of viral vector constructs. By now, there are numerous efficient retrovirus- or lentivirus-based methods which allow stable genomic incorporation of the foreign DNAs with high gene product expression levels. However, virus-based gene therapy technologies also have serious draw-backs, including safety concerns of virus production, and the preferential incorporation of foreign genes into active host gene loci, which may cause uncontrolled proliferation of the gene-modified stem cells [1, 2]. On the other hand, although non-viral gene delivery techniques are usually considered to be less efficient, with the emergence and refinement of the transposon-based

Author contributions: T.I.O.: conception and design; collection and assembly of data; data analysis and interpretation; manuscript writing; Á.A.: conception and design; collection and assembly of data; data analysis and interpretation; manuscript writing; A.N.: collection and assembly of data; N.V.: collection and assembly of data; V.K.: collection and assembly of data; A.S.: collection and assembly of data; K.S.: collection and assembly of data; Z.E.: collection and assembly of data; G.V.: collection and interpretation of data; É.K.: collection of data; L.H.: interpretation of data; K.N.: collection and interpretation of data; E.G.: interpretation of data; C.M.: data analysis; L.M.: interpretation of data; Z. Ivics: data analysis; manuscript writing; Z. Izsvák: data analysis and interpretation of data; B.S.: conception and design; financial support; collection and assembly of data; data analysis and interpretation; manuscript writing; final approval of manuscript; T.I.O. and Á.A.: contributed equally to this article.

Correspondence: Balázs Sarkadi, MD, PhD, Membrane Research Group of the Hungarian Academy of Sciences, Semmelweis University and National Blood Center, Dioszegi u. 64., Budapest, H-1113, Hungary. Telephone: +36-1-372-4353; Fax: +36-1-372-4353; e-mail: sarkadi@biomembrane.hu Received December 17, 2008; accepted for publication February 11, 2009; first published online in STEM CELLS *EXPRESS* February 20, 2009. © AlphaMed Press 1066-5099/2009/$30.00/0 doi: 10.1002/stem.45

methods, they certainly represent a plausible alternative to viral applications [3–6].

Current methods for directed tissue differentiation from various HuES cells often apply endogenous morphogenic proteins or invasive chemicals to obtain the tissue(s) of interest [7–9]. However, the use of such artificial cocktail of drugs may affect natural differentiation and, among other problems, may induce undesired gene expression profiles. In addition, it is often advisable to examine whether tissues derived after such invasive attempts reliably represent naturally occurring cell types [10]. A common way of overcoming this problem is to use a combination of tissue-specific promoters and certain drug resistance genes as selection markers. This method can provide the solution for obtaining the desired cell types, as it is based on "spontaneous" differentiation of HuES cells (see [11]) and selecting for the differentiated cell types in a later phase, e.g., by certain antibiotics [12, 13]. In principle, this could result in the enrichment of desired tissues in large quantities and most likely provide less interference with the gene expression profiles of the tissues of interest. However, it still requires drug selection which may induce an altered way of differentiation, again asking for the need to prove that the obtained tissues represent naturally occurring cell types [14].

The use of tissue specific promoters generates another technical problem, as in this case transgene expression is not detectable in the undifferentiated HuES cells [15]. Especially for non-viral applications, where the efficiency of gene delivery is generally low, this represents a serious disadvantage. To avoid this, one can use another, constitutive promoter to drive the expression of an additional selection marker gene, and first select for transgene delivery, then, after spontaneous differentiation, for specific tissues. Apart from being cumbersome and time-consuming, this also creates another technically challenging problem, requiring either the codelivery of different transgenes or the use of a larger cargo vector, often resulting in lower delivery efficiency.

Delivering transgenes into HuES cells also requires a careful design as to which promoters are to be used for yielding high expression levels. Previous studies are controversial concerning the applicability of widely used transgene promoters. The cytomegalovirus immediate early 1 (CMV) promoter has been claimed to be a useful system for expressing transgenes in undifferentiated mouse embryonic stem (ES) cells [16]. Some studies, however, found that the CMV promoter has only a moderate activity when compared with the elongation factor 1α (EF1α) promoter [17]. Other investigations revealed that the CMV promoter is strongly silenced in ES cells by means of epigenetic modification via DNA methylation [18, 19]. Chung et al. [20] carried out a systematic study in comparing three promoters in undifferentiated and differentiated mouse ES cells. They demonstrated that the EF1α and the CMV-chicken β-actin fusion (a version of CAG-CMV-Actin-Globin) promoters yielded robust gene expression in undifferentiated mouse ES cells and effective transgene expression both in embryoid bodies and differentiated neuronal precursors. On the other hand, the CMV promoter was only active in the neuronal precursors [20]. Similar systematic study was carried out in human ES cells, using a lentiviral transduction method [21]. It was demonstrated that the CAG promoter drove gene expression in the ES cells more efficiently than the (strongly silenced) CMV promoter, whereas the EF1α and the phosphoglycerate kinase (PGK) promoters showed much stronger transgene expression and were thereby claimed to be the promoters of choice for HuES cells. Nevertheless, it is still unsettled which promoter is to be used in ES cells, especially if one needs to have strong transgene expression in certain types of tissues, originated from the genetically modified ES cells.

In this study, we have developed a non-viral methodology for efficient and stable gene delivery into HuES cells. We applied *Sleeping Beauty* (SB) transposition [4, 22] for non-viral gene delivery and expression, examined the efficiency of stable gene delivery, as well as the expression level and the effect of a reporter gene construct on early stem cell differentiation. In this system, enhanced green fluorescent protein (EGFP)-expressing SB transposon-based vectors were delivered into HuES cells together with a hyperactive version of the SB transposase that shows ~32-fold higher activity than the originally reported first-generation transposase [3]. By comparing this system to a lentiviral gene delivery into the HuES cells, we provide a detailed description of the two stable gene modification methods.

In addition, we have developed an efficient and relatively simple genetic system to selectively enrich transgene-expressing cardiomyocytes differentiated from HuES cells. We found that one version of the widely used CAG promoter exhibits a previously unknown "double-feature" characteristic: it can be used to mark genetically modified undifferentiated HuES cells and it also provides a selection platform specifically for cardiomyocytes during subsequent spontaneous differentiation. We documented that this unusual behavior is independent of gene delivery methods, and the sequence, copy number, and chromosomal integration site of the transgene.

Materials and Methods

Embryonic Stem Cell Culturing and Differentiation

The HuES cell lines HUES9 (originally provided by Dr. Douglas Melton, Harvard University) and BG01V (from ATCC) were maintained essentially as described earlier [23]; cells from passage no. 35 were used for these analyses. Differentiation of the HuES cells were initiated spontaneously via the embryoid body (EB) formation pathway [23]. Cell types were identified by morphological signs under phase contrast light microscopy and by immunostaining for various protein markers.

Promoter and Vector Constructs

Plasmids used in this project contained the cDNA of the highly fluorescent marker amaxaGFP or EGFP (Fig. 1A; the original CAG-amaxaGFP transposon plasmid was kindly provided by Evelyn Zeira via collaboration in the EU FP6 INTHER project). By a Hind-III and NheI restriction digestion followed by ligation, we cloned the CMV viral promoter, the CAG promoter, the human PGK promoter or the short version of the human EF1α promoter upstream of the transgene in the transposon vectors (Fig. 1B). For the lentiviral constructs, the used transgene cassettes were exactly the same, as in all cases the entire transcription unit was removed from the transposon vector by restriction digestion and inserted into the viral vector by blunt end ligation (Fig. 1C).

Transposon-Based Gene Delivery

For the SB transposase, we used an enhanced version of the enzyme having 32 times higher activity than the originally reconstructed transposase ([3] and unpublished results); for transposition control, we applied a DDE inactive mutant form of the enzyme [22]. For transfection, the FuGENE® 6 (Roche Applied Science, Rotkreuz, Switzerland, http://www.roche-applied-science.com) reagent was used according to the manufacturer's instruction. Before transfection, HuES cells were

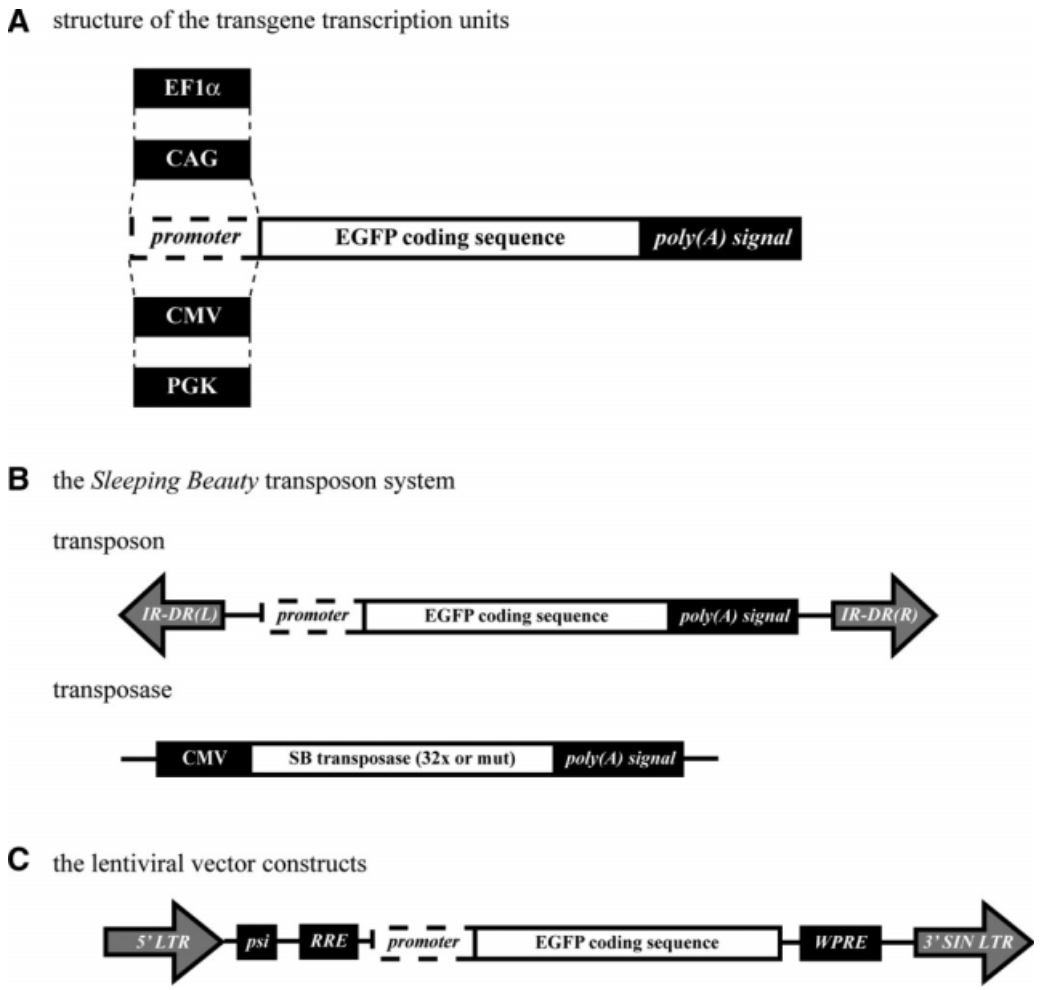

Figure 1. Vector constructs. **(A):** Four different promoters were used to drive transgene expression. **(B):** Structure of the cotransfected plasmids for the *Sleeping Beauty* transposon system. IR-DR(L) and (R) stand for left and right inverted repeat–direct repeat transposon sequence. **(C):** Structure of the lentiviral constructs. Psi and RRE are elements required for viral packaging. Sizes of the different elements are not drawn to scale. Abbreviations: CAG, CMV enhancer-chicken β-actin-rabbit β1-globin fusion; CMV, cytomegalovirus immediate early 1; EF1α, elongation factor 1α; EGFP, enhanced green fluorescent protein; IR-DR(L), inverted repeat-direct repeat, left; IR-DR(R), inverted repeat-direct repeat, right; LTR, (viral) long terminal repeat; PGK, phosphoglycerate kinase; SIN, self-inactivating (virus); SB, Sleeping Beauty; WPRE, woodchuck hepatitis B virus posttranscriptional regulatory element.

separated from the mouse feeder cells and placed on gelatin-coated plates. On the subsequent day, the cells were cotransfected with transposon and transposase plasmids in a 10:1 ratio to avoid overproduction inhibition of the transposase [4, 5]. Next day, the transfected HuES cells were placed back onto the mouse feeder layer to keep them in an undifferentiated state. Although some differentiation could start during the 2 days without feeder cells, in our hands it did not significantly jeopardize the preservation of pluripotency (The current transfection protocol was not efficiently working in a Matrigel system.) To provide evidence that the transposon constructs are capable of transposition, a nested excision PCR method was applied, as described previously. This method amplifies the "footprint" sequence left in the plasmid after transposon excision, whereas no PCR product is obtained if no excision occurs [24]. To determine the integration sites of the transgenes in human genomic DNA, splinkerette PCR and inverse PCR methods were applied, essentially as described earlier [22, 25]. Copy number determination was carried out by real-time PCR using the relative standard curve method with a TaqMan® assay designed specifically for the right inverted repeat sequences of the SB transposon vector. As a single-copy gene control, a predeveloped TaqMan® assays for the RNA subunit of the RNaseP enzyme, was purchased from the manufacturer; analyses were carried out using the StepOne Real-Time PCR System (Applied Biosystems, Foster City, CA, http://www.appliedbiosystems.com).

Lentiviral Transduction Procedure

For the viral-based gene delivery, the third generation lentiviral vector system was used, as described in [26] (Fig. 1C). Determination of virus titers and transduction procedures were performed essentially as described previously [27]. The HUES cells were transduced by a MOI of 2–5, and the further handling of undifferentiated cells and cell differentiation protocols were the same as for the transposon-based gene delivery system.

Flow Cytometry

EGFP-expressing undifferentiated HuES cell colonies were manually selected by using a fluorescent microscope under sterile conditions, to enrich the marker gene expressing population. Such heterogeneous colonies (Fig. 2A) were harvested from the mouse feeder cells, washed with PBS and were

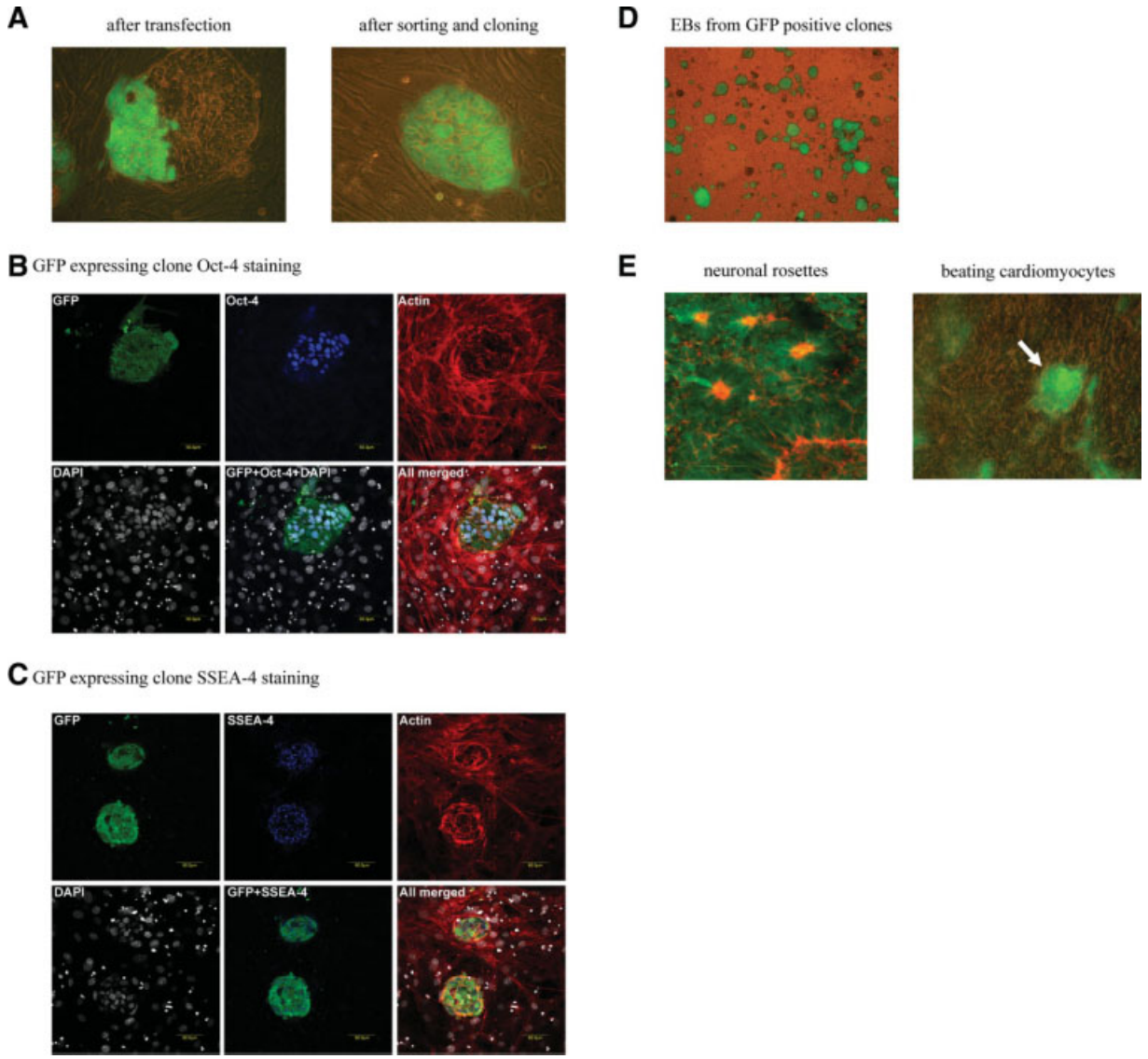

Figure 2. Human embryonic stem cells stably expressing the transgene, exemplified by the CAG-amaxaGFP transgene, delivered by the SB transposon system. **(A):** Morphology and GFP expression of human embryonic stem cell (HUES9) clumps. GFP expression was detected by fluorescence microscopy following transfection (left panel) or after subsequent sorting and cloning of the transgene expressing cells (right panel, magnification: ×200). **(B):** GFP expressing HUES9 clone on mouse feeder cells stained for the Oct-4 embryonic stem cell marker. This transcription factor is localized in the nucleus and it is only expressed in the cells of the HuES clump and not in the surrounding feeder cells. Green: amaxaGFP; blue: Oct-4; red: TRITC-phalloidin, representing actin filaments; white: DAPI staining for the nucleus. Confocal images, scale bars represent 50 μm. **(C):** GFP expressing HUES9 clone on mouse feeders stained for the SSEA-4 embryonic stem cell marker. This plasma membrane protein could only be detected in HuES clump cells and not in the surrounding embryonic feeders. Green: amaxaGFP; blue: SSEA-4; red: TRITC-phalloidin, representing actin filaments; white: DAPI staining for the nucleus. Confocal images, scale bars represent 50 μm. **(D):** Six-days-old EBs formed from GFP expressing HUES9 clones observed by fluorescence microscopy (magnification: ×40). **(E):** Neuronal and myocardial cells differentiated from GFP expressing HUES9 clones. The left panel represents a confocal image of neuronal rosettes (scale bar represents 50 μm). The right panel depicts a typical area with beating cardiomyocytes (white arrow), indicating a prominent difference of amaxaGFP expression between the cardiomyocytes and the other surrounding tissues (phase contrast fluorescence microscopy image, ×40 magnification). Green: GFP fluorescence, red: TRITC-phalloidin, representing actin filaments. Abbreviations: DAPI, 4′,6-diamidino-2-phenylindole; EB, embryoid body; GFP, green fluorescent protein.

sorted based on GFP fluorescence using the fluorescence-based cell sorting (FACS) Aria High Speed Cell Sorter (Beckton-Dickinson, San Jose, CA, http://www.bdbiosciences.com). Mock-transfected HuES cells were measured to set the level for EGFP-positivity with the FACS Diva analysis software; propidium iodide staining was used to gate out the nonviable cells. Sorted EGFP positive single cells were placed onto the mouse feeder cells and monitored until the formation of surviving clones. For further gene expression analysis, differentiated cells from EGFP-expressing HuES clones were sorted into four different artificial fractions based on increasing EGFP fluorescent signal intensity. Cells obtained from different fractions were washed with PBS and immediately resuspended in Trizol (Invitrogen, Carlsbad, CA, http://www.invitrogen.com) for further RNA analysis.

Immunohistochemical Assays

Immunostaining procedures were performed essentially as described previously [23]. For stem cell markers, monoclonal antibodies against Oct-4 (Santa Cruz Biotechnology,

Santa Cruz, CA, http://www.scbt.com), SSEA-4, or podocalyxin (R&D Systems Inc., Minneapolis, MN, http://www.rndsystems.com) were used. TRITC-phalloidin (Sigma-Aldrich, St. Louis, MO, http://www.sigmaaldrich.com) staining was used for the detection of actin filaments, anti-myosin IIA (Santa Cruz Biotechnology) was used for nonmuscle cell detection, DAPI (Invitrogen) was used for nuclear staining, and GFP was evaluated by direct fluorescence. All documented measurements were carried out on cells from four independent GFP-expressing clones, at least in triplicate stainings.

RNA Analysis

RNA isolation was carried out from cells collected in Trizol reagent (Invitrogen) according to the manufacturer's instruction. cDNA samples were prepared from 0.1 μg total RNA using the Promega Reverse Transcription System Kit, as specified by the manufacturer. Tissue specific genes were selected using the TiProD tissue-specific promoter database (tiprod.cbi.pku.edu.cn:8080/index.html). The following markers were selected: OCT-4 and NANOG transcripts as undifferentiated stem cell markers; ACTC (cardiac-specific α-actin), NPPA (natriuretic peptide precursor A) and PLN (phospholamban) genes as cardiac specific markers; PAX6 (paired box gene six) gene as an early marker for neuronal differentiation; CAPG (capping protein (actin filament), gelsolin-like) gene as a skin differentiation marker; P0 ribosomal protein as endogenous controls. Predeveloped real-time TaqMan® assays for the listed genes were purchased from Applied Biosystems. For quantifying EFGP mRNA, specific TaqMan® assay were designed for the cDNA sequence. Real-time PCR analyses were carried out using the StepOne Real-Time PCR System (Applied Biosystems), according to the manufacturer's instructions.

RESULTS

Comparison of the Transposon-Based and Lentiviral Gene Delivery Systems

An important goal in ES cell biology is to create transgenic stem cells with stable transgene expression to provide a source for potentially any type of tissues. The first difficulty to overcome during this process is to achieve a high rate of stable gene delivery into HuES cells. For this purpose, we applied the SB transposon system and tested several constitutive promoters previously shown to drive marker gene expression in ES cell lines, namely, the cytomegalovirus immediate early 1 (CMV) promoter, the EF1α promoter, the PGK promoter, and the artificial CMV enhancer—chicken β-actin–rabbit β1-globin fusion (CAG) promoter. To compare the efficiency and the applicability of viral and non-viral gene delivery, we have also applied a lentiviral-based vector system to express the EGFP marker gene with the promoters mentioned earlier (Fig. 1).

In the case of the transposon system, the lipid-based transfection was less efficient (5%–10%) than viral transduction (in most cases above 20%), but the selected transgene expressing cells were found to be stable and useful for later cloning and differentiation studies in both cases. As expected from the data in the literature, the CMV promoter was not applicable in our HuES cell lines, as transgene expression rapidly faded away soon after transfection/transduction, most likely due to the silencing of this viral promoter [18, 19]. The other constitutive promoters (PGK, EF1α, and CAG) expressed the transgene in undifferentiated and in differentiating HuES cells. To avoid the problem of copy number variation when comparing expression levels among different tissue types after differentiation, we established clones from transgenic stem cells expressing EGFP by all these three different promoters. For each case, a manual enrichment of the EGFP-expressing HuES cells was followed by a flow cytometry sorting and the selected cells were placed back onto mouse feeders at a density to allow individual clones to grow separately (Fig. 2A). These clones preserved their ES cell characteristics, as they expressed the ES cell markers Oct-4, SSEA-4, and podocalyxin (Fig. 2B, 2C, and data not shown). The uniform EGFP expression in these clones, grown in an undifferentiated state for 3–20 passages, was preserved; thus, at least in this period, a stability of foreign gene expression was observed (supporting information Fig. 2A).

Subsequently, we followed differentiation of these HuES clones by embryoid body (EB) formation in vitro. The resulting EBs and the various tissue types differentiated from them were similar to those observed in the untransfected HuES cells, and there was no observable change in the general pattern or timing of these differentiating clones (Figs. 2D, 2E). Concerning cardiomyocytes, 19–33 beating areas/100 plated EBs were detected in different transgenic clones which was not significantly different from the average number of 28 beating areas observed during differentiation of control HuES cells. A detailed analysis of the gene expression patterns in these differentiated colonies is underway in our laboratory (see discussion). Collectively, these experiments indicated that neither the transduction nor the combined transfection/transposition procedure altered the cardiomyocyte differentiation potential of HuES cells per se and stable transgene expression could be achieved in these pluripotent cells by our gene delivery methods.

Robust CAG Promoter-Driven Gene Expression in Cardiomyocytes

HuES cells form EBs when detached from the plate surface, and under well-defined conditions the EBs spontaneously differentiated into various tissue types (such as neuronal, epithelial, myocardial, and other progenitors [7, 23, 28, 29]). Cardiomyocytes that differentiate from HuES cells can easily be recognized by the spontaneous contractions of small tissue patches, by the immunochemical recognition of cardiac-specific markers [30], and by a clear response to pharmacological modulations (increased contraction frequency after adrenaline, decreased frequency, and arrest after calcium channel blockers [31, 32]). Although we do not include a detailed documentation here, the cardiomyocytes differentiated from the parental or gene-modified HuES cells showed all these parameters, without observable differences caused by the actual methods of gene transfer (supporting information Video 1). As the BG01V cell line has a lower potential to generate beating cardiomyocytes, most of the detailed experiments were carried out in the HUES9 cell line.

When comparing various tissue types differentiated from such EGFP-expressing HuES clones, an intriguing observation in CAG promoter-driven EGFP cells was that the fluorescent signal intensity in cardiomyocytes was always significantly higher than in the other surrounding tissue types. As the phenomenon was first seen for SB transposon-derived clones, this gene delivery experiment was repeated several times and the outcome was always the same: CAG-driven EGFP expression was extremely strong in differentiated cardiomyocytes as compared with cells with PGK promoter- or EF1α promoter-driven transgenes (Figs. 2E and 3). In fact, in the case of the CAG promoter, all beating cardiomyocytes could easily be

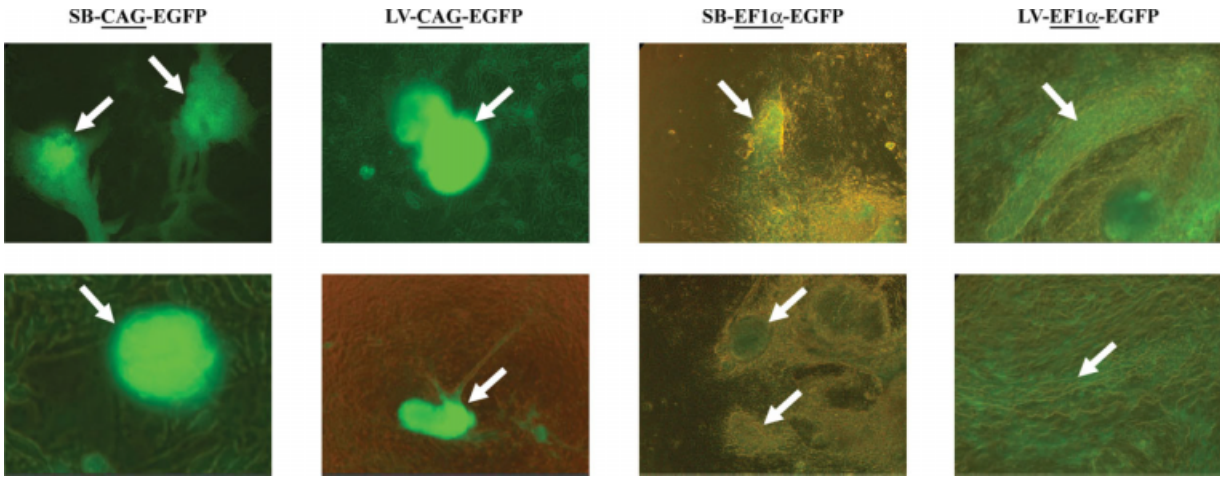

Figure 3. The "double-feature" phenomenon of the CAG promoter is independent of the gene delivery method. Several independent differentiations from transposon (SB) or lentiviral (LV) clones expressing EGFP resulted in numerous spontaneously contracting cardiomyocytes (white arrows, see also supporting information Videos 2 and 3). When the CAG promoter was used, the EGFP fluorescence in cardiomyocytes was always extremely high so that it masked the green fluorescence in the surrounding tissues (left "CAG" panels). In cells expressing the EF1α-driven transgene, cells were expressing the EGFP almost uniformly and did not show such an expression bias toward cardiomyocytes (right "EF1α" panels). Abbreviations: CAG, CMV enhancer-chicken β-actin-rabbit β1-globin fusion; EF1α, elongation factor 1α; EGFP, enhanced green fluorescent protein; LV, lentiviral; SB, Sleeping Beauty.

identified by the extremely high transgene expression. This unexpected behavior of a constitutive promoter allowed solving two problems at the same time: following transgene expression in undifferentiated cell types and selecting for a particular tissue type after differentiation.

The "Double-Feature" Behavior of the CAG Promoter Is Independent of the Gene Delivery Method and the Sequence, Copy Number, and the Integration Sites of the Transgenes

To prove that the phenomenon is a characteristic of the CAG promoter, we first compared the results of two different CAG-driven fluorescent transgenes (EGFP and amaxaGFP) delivered into the HuES cells by the SB transposon system. EGFP (e.g., www.clontech.com) and the amaxaGFP (www.amaxa.com) are sequentially unrelated proteins, isolated from different organisms. Despite this sequence difference, the same phenomenon was observed, that is the differentiated cardiomyocytes showed a markedly higher fluorescent signal as compared with other differentiated cell types (Figs. 2E, 3, and supporting information Video 2, 3). Lentiviral gene delivery provided an additional proof that the CAG promoter, but not the EF1α or the PGK promoters, has this behavior of being transcriptionally extremely active selectively in cardiomyocytes (Fig. 3 and data not shown). Thus, this feature of the CAG promoter is independent of the gene delivery method.

To further characterize the phenomenon, we compared clones with different copy numbers of the CAG-EGFP transgene for both transposon-derived and lentiviral-derived HuES clones. The transgene copy numbers determined by real-time PCR were found to be 3–6 in our experiments. The differentiation-linked EGFP expression behavior was detected in all cases, clearly showing that copy number differences do not influence the "double-feature" behavior.

To document the independence of this phenomenon from transgene integration into specific host loci, we carried out integration site determination for transposon-based amaxaGFP-expressing HuES cells (Fig. 4). We identified several different genetic loci (being on different chromosomes) in which SB transposition occurred. In all cases, although these sites were apparently not related to any cardiomyocyte-related protein coding genes or miRNA clusters, an increased expression of EGFP in the cardiomyocytes was observed. All these data indicated that it is not the site of integration, but the CAG promoter itself that is responsible for the observed effect.

Selection for High, CAG Promoter-Mediated Transcription Enriches Differentiated Cells Expressing Cardiomyocyte Markers

In principle, higher EGFP fluorescence in particular cell types can be the result of various factors, including higher transcription or translation rate, or certain posttranslational modifications. Alternatively, the size or the higher cytoplasm/nucleus ratio in the observed cell types could also account for the higher signal intensity. To distinguish among these possibilities, we separated cells based on EGFP fluorescence by FACS at the 30th day of differentiation and isolated total cellular RNA from each separated fraction. Four different EGFP fractions were analyzed from cells differentiated from either CAG-EGFP or EF1α-EGFP-expressing HuES cell clones, created by either lentiviral-based or transposon-based gene delivery (Fig. 5B and supporting information Fig. 1B). By real-time quantitative PCR analysis, we measured the transcriptional profiles of cardiac-specific, early neuron-specific, and skin-specific marker genes, as well as stem cell-specific (pluripotency marker) genes, and also the EGFP mRNA levels in the different cell populations. As expected, the expression levels of pluripotency marker genes, OCT-4 and NANOG, were high in the undifferentiated HuES clones, whereas practically undetectable in the differentiated populations (Fig. 5A and supporting information Fig. 1A). In the differentiated cells, the EGFP transcription level closely correlated with the fluorescent signal intensity, thereby proving that high EGFP intensity was the result of increased transcription, as a consequence of increased promoter activity (Fig. 5C and supporting information Fig. 1C).

When analyzing cardiac-specific transcripts, we measured the mRNA levels of the ACTC, PLN, and NPPA genes (see Materials and Methods). For all cells originated from CAG-EGFP clones (irrespective of the gene delivery method applied), these cardiac-specific transcripts exhibited a

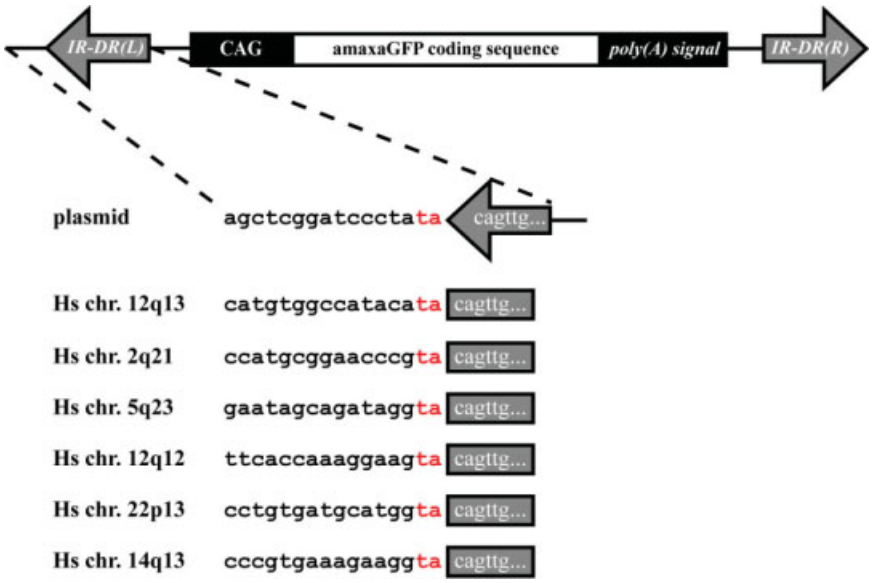

Figure 4. Examples of transposon integration sites. Splinkerette and inverse PCR techniques were used to determine the SB-CAG-amaxaGFP integration sites in HuES clones. In all cases, a "ta" sequence (marked by red) was found in the genomic locus next to the transposon IR-DR sequence which is a footprint of *Sleeping Beauty* transposition [25]. None of the detected integration sites were apparently related to any cardiomyocyte-associated protein coding genes or miRNA clusters.

significant, several-fold higher level in high EGFP expressing populations, indicating that in differentiated cells, cardiomyocytes were located in the fractions with high EGFP signal intensity (as observed earlier by fluorescence and confocal microscopy). On the other hand, in fractions from cells expressing the EF1α-EGFP transgene there was no such correlation: the examined cardiac specific genes showed a random distribution among the different EGFP expressing populations (Fig. 5C and supporting information Fig. 1C). When analyzing early neuron-specific (PAX6) or skin-specific (CAPG) transcripts, we did not find any correlation between their mRNA levels and EGFP transcriptional levels, and these genes had random expression profiles among different fractions for both CAG-EGFP- and EF1α-EGFP-expressing cells (Fig. 5C and supporting information Fig. 1C). Taken together, these expression studies provide evidence that the transcriptional activity of the CAG promoter, in contrast to other constitutive promoters examined, is significantly higher, selectively in cardiomyocytes. In addition, it also raises the possibility of flow cytometric separation of cardiomyocytes, based on high CAG promoter-driven transgene expression, especially because the currently available cardiac-specific cell surface antibodies require fixation and/or permeabilization of the cells (supporting information Fig. 2B).

Deciphering the "Double-Feature" Behavior by Bioinformatics Approaches

The artificial CAG promoter used in this study contains the CMV enhancer region, two sequences from the chicken β-actin promoter and one short part of the rabbit β1-globin promoter (Fig. 6A). To understand the sequential reasons behind the "double-feature" phenomenon, we conducted a bioinformatics analysis of the CAG promoter for potential tissue-specific transcription factor binding sites, by using the program MatInspector 7.7.3. When the analysis was narrowed down to heart-specific transcription factors, several potential binding sites were revealed (Fig. 6B). An especially intriguing feature of the CAG promoter is that it contains a consensus Nkx2.5 binding site (nt 234–248), three serum response factor (SRF.01, SRF.02, and SRF.03) binding sites (nt 317–335, 55–73, 595–613), as well as a MEF2-SL1.01 binding site (nt 624–646). These are all located within the CMV enhancer and the first chicken β-actin promoter region (Fig. 6B).

Several studies have documented that Nkx2.5 expression is specifically required for cardiogenesis, and the expression of this transcription factor is present from the early, neonatal development, to the fully matured form of cardiomyocytes [35–38]. SRF proteins were documented to cooperate with MEF2 and Nkx2.5 to promote cardiogenesis, and the activation of SRF by its interaction with myocardin provides a strict tissue-specificity for SRF binding promoters in the cardiac tissues [39]. Based on these data, it is not unexpected that the CAG promoter, most probably through multiple interactions with cardiac-specific transcription factors, shows an exquisite overexpression in cardiomyocytes.

Discussion

To establish stable transgene integration into ES cells, the careful design of the gene delivery method and the choice of promoters are crucial points. Among non-viral gene deliveries, transposon-based methods are becoming increasing popular, because their advantages outweigh their potential technological drawbacks. They certainly require less sophisticated and expensive facilities, their use is simpler than viral applications, and their use can also overcome safety and ethical issues currently associated with viral-based vector systems. One of the most effective transposon-based platforms is the SB system. One particular advantage of the SB system in the context of potential human applications is undoubtedly the random genomic integration profile, which lowers the chance of insertional mutagenesis or induced oncogenesis. These problems represent a serious drawback for viral-based gene delivery methods [1, 2]. A further advantage of the SB system is the availability of hyperactive transposases, providing highly efficient gene delivery tools [3, 24]. There is one report in which the SB system was successfully applied in HuES cells [33], but this study used an older version of the transposase (SB11), and a nucleofection technology in a given HuES cell line, which may not be directly applicable for other stem cell cultures. Therefore, we aimed to examine the use of a hyperactive SB transposase, combined with a lipid-based transfection methodology, and compared this with the efficient lentiviral gene delivery system.

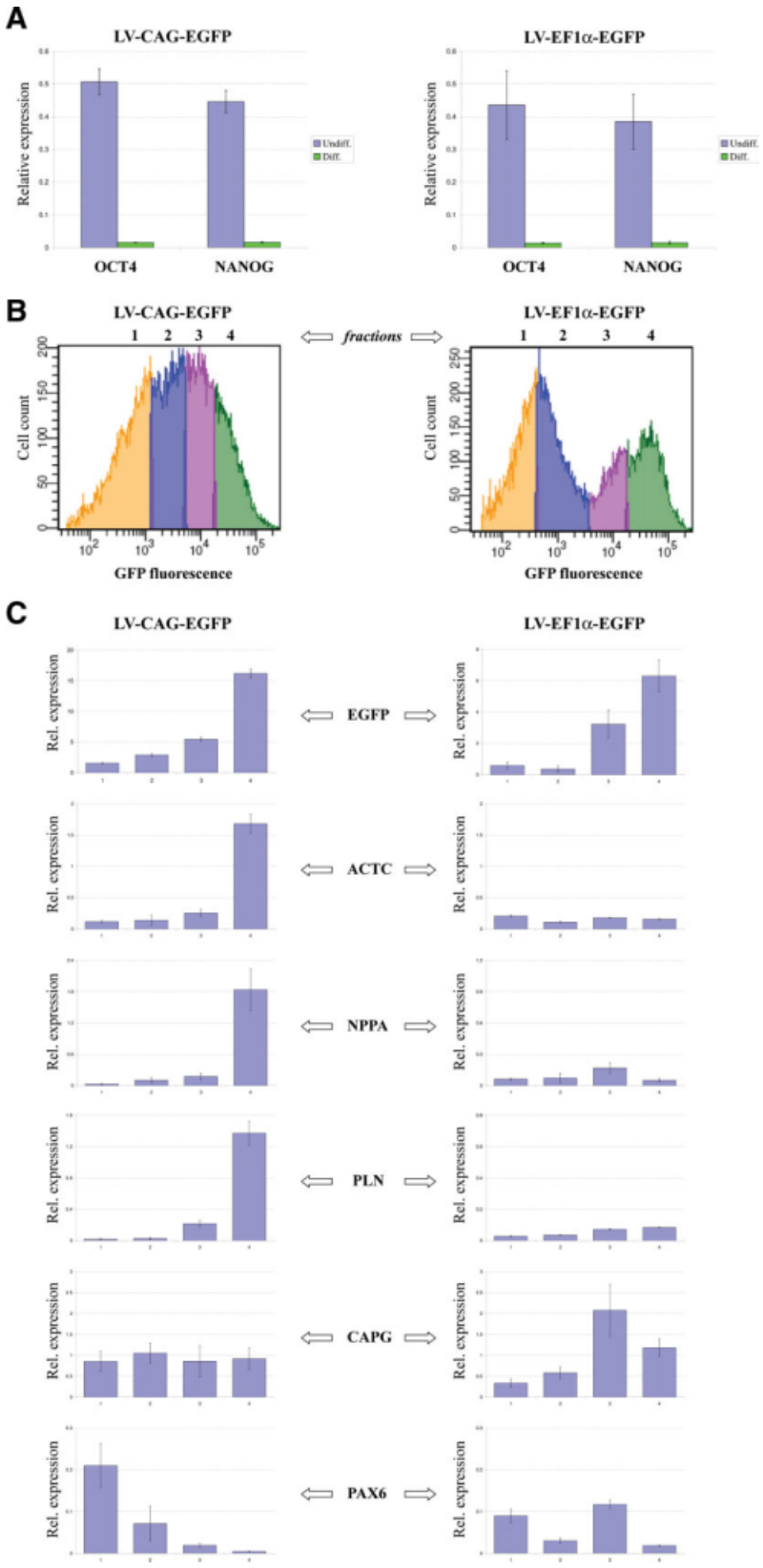

Figure 5. Expression profiles of HuES clones expressing either CAG-EGFP or EF1α-EGFP transgenes after lentiviral gene delivery. **(A):** Comparing the expression levels of stem cell markers in undifferentiated (Undiff.) and in differentiated (Diff.) HuES cells. The transcription levels of OCT-4 and NANOG were almost undetectable and at least 40-times lower in differentiated cells. **(B):** Fluorescence-based cell sorting diagrams indicating the sorted populations of differentiated EGFP-expressing cells. For unknown reasons, two major EGFP peaks were detected in this particular experiment after differentiating EF1α-EGFP expressing HuES clones. Total RNAs were isolated from cells collected from the four indicated fractions in both cases, representing cells with increasing EGFP signal intensity. **(C):** Transcription levels of EGFP and differentiation specific endogenous genes in the four different fractions. Increasing EGFP transcript level positively correlated with increasing fluorescence signal intensity. mRNA levels of cardiac specific genes (ACTC, NPPA, and PLN) showed a several fold increase in the cell population with the highest EGFP intensity for the CAG promoter-driven transgene, whereas no such tendency was observed for the EIF1α promoter-driven transgene. No correlation between EGFP transcription level and skin-specific (CAPG) or early neuron-specific (PAX6) gene transcripts could be identified for either of the promoters. In all real-time PCR experiments, results represent mean values ± S.E.M. of at least three independent measurements; transcription levels were normalized to the ribosomal P0 endogenous control. Abbreviations: ACTC, cardiac-specific alpha-actin; CAG, CMV enhancer-chicken β-actin-rabbit β1-globin fusion; CAPG, capping protein (actin filament), gelsolin-like; EF1α, elongation factor 1α; (E)GFP, (enhanced) green fluorescent protein; LV, lentiviral; NPPA, natriuretic peptide precursor A; PLN, phospholamban.

Recent, as yet unpublished experiments clearly indicate the higher specific integration efficiency of the hyperactive transposase system in tissue-derived stem cells (Ivics and Izsvak, personal communication). Our experiments demonstrate that the same SB transposon system is an effective gene delivery method that can be applied to create stable cell lines for a given transgene even if the transfection efficiency is low and a drug-based selection method is not applied. These

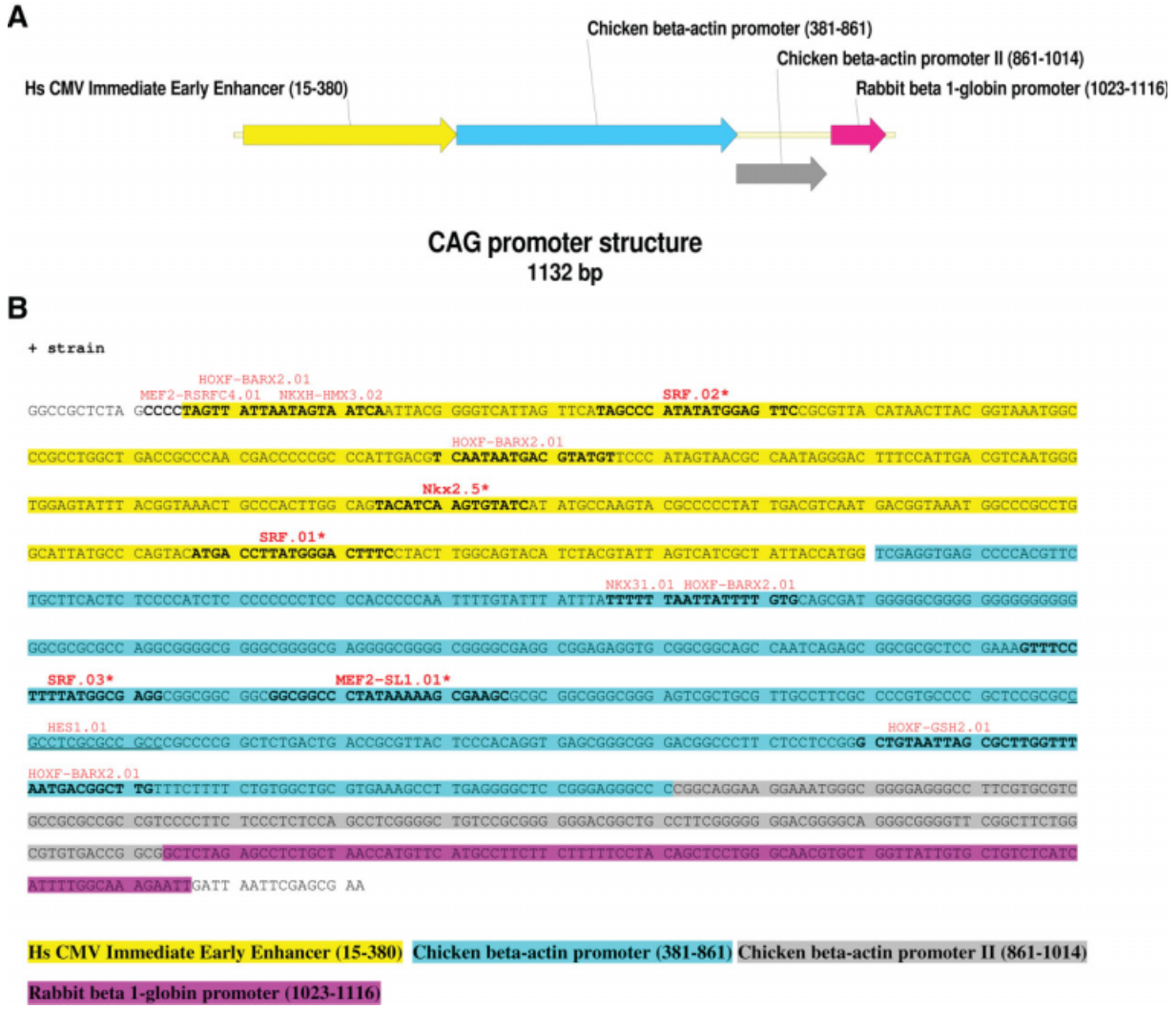

Figure 6. Structure of the applied CAG promoter. **(A):** Proportional representation of the distinct elements of the fusion promoter. **(B):** The predicted cardiac-specific transcription factor binding sites. The most prominent one is the complete consensus site of the Nkx2.5 homeodomain factor, located in the CMV Immediate Early Enhancer element (nt 234–248). See text for further details. Color coding is intentionally comparable between the two panels. Abbreviations: CAG, CMV enhancer-chicken β-actin-rabbit β1-globin fusion; CMV, cytomegalovirus.

conditions especially apply to HuES cells, which are difficult subjects for transfection, and may not be selectable by commonly used antibiotics without jeopardizing their pluripotency status (Fig. 2). We propose that the transposon protocol presented here could be a method of choice for gene delivery into HuES cells.

To assure the similarity of the transgene-expressing HuES cells and their differentiation to the normal parental cells, we have analyzed a number of gene expression patterns, including the expression of the ABCG2 protein [23], and found no apparent differences in these parameters (Figs. 2, 5 and data to be reported elsewhere). For a systematic comparison of the differentiation patterns in gene-modified cells, a detailed analysis of a full gene expression profile by a TaqMan based real-time quantitative PCR assay system is underway in our laboratory.

The use of suitable promoter(s) for foreign gene expression in stem cells is still an unresolved issue, although several studies described the potential silencing or the variable transcriptional efficiency of commonly used promoters [19–21]. It is widely accepted, however, that some housekeeping promoters, such as the PGK or the EF1α promoters, as well as the constitutive artificial CAG promoter most likely represent the best choices for stem cell applications, as these promoters also remain active in a wide variety of differentiated tissues, although with variable intensities [19–21, 34]. On the other hand, the need to express transgenes specifically in a given tissue raises the problem of controlling transgene delivery and tissue-specific transcription at the same time. Current solutions to this problem offer the use of two different transcription units separately (with the problem of the delivery of two plasmids or viruses), or in one cargo (with the problem of larger plasmid or virus vector size), both of which decrease delivery efficiency. Based on the present study, we propose that the use of the "double-feature" CAG promoter solves both problems: it provides a means for transgene detection in undifferentiated cells, as well as a later separation platform for cardiomyocytes, based on its selectively high transcription rate in these cell types (Figs. 3 and 5). We provided evidence that this feature is independent of the gene delivery method, transgene sequence, copy number, or integration sites (Figs. 3 and 4). We are currently investigating the applicability of cardiomyocytes selected by this method for potential therapeutic purposes. As an important first criterion, these cells

show negligible OCT-4 and NANOG expression (Fig. 5A), indicating the absence of undifferentiated cells that might represent a risk of tumor formation [30–32].

It should be noted that the term "CAG promoter" is inconsistently used in the scientific literature. The version applied in our experiments is probably very close to the originally created artificial promoter (Fig. 6A), as it contains the three crucial elements (*C*MV enhancer, *β-a*ctin sequence, rabbit *β*1-*g*lobin sequence), although it has two shorter parts of chicken actin promoter rather than the original long sequence [40]. It is often impossible to decipher the exact "CAG" sequence used by a particular research group, but from the size of the construct in several cases it is likely different from the "canonical" promoter. A common mistake is that a construct without the rabbit *β*1-*g*lobin sequence is also called a "CAG" promoter [41, 42], or the CMV enhancer sequence is sometimes shortened, e.g., in the commercially available pDRIVE-CAG plasmid (www.invivogen.com). If our bioinformatics analysis correctly identifies the Nkx2.5 and the serum response element binding sites as responsible sequences for the cardiac-specific upregulation, than many of the generally used "CAG" promoters may not show the "double-feature" characteristics. Currently, we have embarked on a detailed experimental analysis of these questions; deletions and binding factor consensus site mutations are hoped to reveal the background of this phenomenon. Moreover, as this phenomenon holds true for two human-derived ES cell lines, HUES9 and BG01V, we have started similar experiments in mouse ES cells, and preliminary results indicate that the same CAG promoter also drives cardiac-specific marker expression in this species (data not shown).

An important issue is whether it is possible to design "double-feature" promoters with different tissue specificities. By understanding the exact sequence characteristics from the current constellation of the CAG promoter, we hope to be able to create other "double-feature" promoters. These could become important tools in ES cell research and in future gene therapy applications.

Conclusion

In this study, we show that hyperactive SB transposons represent a powerful gene delivery methodology in HuES cell-based applications. By random transgene integration, this method provides an attractive alternative of virus-based applications. We also document that the CAG promoter shows unexpected characteristics when applied in HuES cells. This promoter has a long-term activity in undifferentiated cells, thus enabling the investigator to control the efficiency of transgene delivery and selecting for transgene-expressing cells. At the same time, the CAG promoter also provides a possibility in later differentiation stages for cardiomyocyte selection, via its unique, tissue-specific transcriptional upregulation. The application of this "double-feature" promoter provides a tool for human cardiomyocyte recognition and separation, and the system may be developed to follow various cell differentiation pathways.

Acknowledgments

The authors are grateful to Zsuzsanna Sebestyén for the excellent technical assistance. The authors appreciate the gift of HUES9 cell line by Dr. Douglas Melton, HHMI, and the original CAG-amaxaGFP transposon plasmid by Evelyn Zeira, Hadassah Medical Organization, Israel. This work was supported by EU FP6-INTHER (LSHB-CT-2005018961), OTKA (AT 048986, and NK72057), NKFP-1A-060/2004, ETT 405/2006, and KKK grants.

Disclosure of Potential Conflicts of Interest

The authors indicate no potential conflicts of interest.

References

1 Schroder AR, Shinn P, Chen H et al. HIV-1 integration in the human genome favors active genes and local hotspots. Cell 2002;110:521–529.
2 VandenDriessche T, Collen D, Chuah MK. Biosafety of onco-retroviral vectors. Curr Gene Ther 2003;3:501–515.
3 Ivics Z, Izsvak Z. Transposons for gene therapy! Curr Gene Ther 2006;6:593–607.
4 Izsvak Z, Ivics Z. Sleeping beauty transposition: Biology and applications for molecular therapy. Mol Ther 2004;9:147–156.
5 Izsvak Z, Ivics Z, Plasterk RH. Sleeping Beauty, a wide host-range transposon vector for genetic transformation in vertebrates. J Mol Biol 2000;302:93–102.
6 Ivics Z, Izsvak Z. Transposable elements for transgenesis and insertional mutagenesis in vertebrates: A contemporary review of experimental strategies. Methods Mol Biol 2004;260:255–276.
7 Lev S, Kehat I, Gepstein L. Differentiation pathways in human embryonic stem cell-derived cardiomyocytes. Ann N Y Acad Sci 2005; 1047:50–65.
8 Karner E, Unger C, Sloan AJ et al. Bone matrix formation in osteogenic cultures derived from human embryonic stem cells in vitro. Stem Cells Dev 2007;16:39–52.
9 Chen K, Wu L, Wang ZZ. Extrinsic regulation of cardiomyocyte differentiation of embryonic stem cells. J Cell Biochem 2008;104:119–128.
10 Tomescot A, Leschik J, Bellamy V et al. Differentiation in vivo of cardiac committed human embryonic stem cells in postmyocardial infarcted rats. Stem Cells 2007;25:2200–2205.
11 Passier R, Oostwaard DW, Snapper J et al. Increased cardiomyocyte differentiation from human embryonic stem cells in serum-free cultures. Stem Cells 2005;23:772–780.
12 Kolossov E, Bostani T, Roell W et al. Engraftment of engineered ES cell-derived cardiomyocytes but not BM cells restores contractile function to the infarcted myocardium. J Exp Med 2006;203: 2315–2327.
13 Huber I, Itzhaki I, Caspi O et al. Identification and selection of cardiomyocytes during human embryonic stem cell differentiation. FASEB J 2007;21:2551–2563.
14 Duan Y, Catana A, Meng Y et al. Differentiation and enrichment of hepatocyte-like cells from human embryonic stem cells in vitro and in vivo. Stem Cells 2007;25:3058–3068.
15 Gallo P, Grimaldi S, Latronico MV et al. A lentiviral vector with a short troponin-I promoter for tracking cardiomyocyte differentiation of human embryonic stem cells. Gene Ther 2008;15:161–170.
16 Ward CM, Stern PL. The human cytomegalovirus immediate-early promoter is transcriptionally active in undifferentiated mouse embryonic stem cells. Stem Cells 2002;20:472–475.
17 Zeng X, Chen J, Sanchez JF et al. Stable expression of hrGFP by mouse embryonic stem cells: Promoter activity in the undifferentiated state and during dopaminergic neural differentiation. Stem Cells 2003; 21:647–653.
18 Krishnan M, Park JM, Cao F et al. Effects of epigenetic modulation on reporter gene expression: Implications for stem cell imaging. FASEB J 2006;20:106–108.
19 Liew CG, Draper JS, Walsh J et al. Transient and stable transgene expression in human embryonic stem cells. Stem Cells 2007;25: 1521–1528.
20 Chung S, Andersson T, Sonntag KC et al. Analysis of different promoter systems for efficient transgene expression in mouse embryonic stem cell lines. Stem Cells 2002;20:139–145.
21 Xia X, Zhang Y, Zieth CR et al. Transgenes delivered by lentiviral vector are suppressed in human embryonic stem cells in a promoter-dependent manner. Stem Cells Dev 2007;16:167–176.

22 Ivics Z, Hackett PB, Plasterk RH et al. Molecular reconstruction of Sleeping Beauty, a Tc1-like transposon from fish, and its transposition in human cells. Cell 1997;91:501–510.
23 Apati A, Orban TI, Varga N et al. High level functional expression of the ABCG2 multidrug transporter in undifferentiated human embryonic stem cells. Biochim Biophys Acta 2008;1778:2700–2709.
24 Ivics Z, Katzer A, Stuwe EE et al. Targeted Sleeping Beauty transposition in human cells. Mol Ther 2007;15:1137–1144.
25 Vigdal TJ, Kaufman CD, Izsvak Z et al. Common physical properties of DNA affecting target site selection of sleeping beauty and other Tc1/mariner transposable elements. J Mol Biol 2002;323:441–452.
26 Schambach A, Galla M, Maetzig T et al. Improving transcriptional termination of self-inactivating γ-retroviral and lentiviral vectors. Mol Ther 2007;15:1167–1173.
27 Ujhelly O, Ozvegy C, Varady G et al. Application of a human multidrug transporter (ABCG2) variant as selectable marker in gene transfer to progenitor cells. Hum Gene Ther 2003;14:403–412.
28 Schulz TC, Palmarini GM, Noggle SA et al. Directed neuronal differentiation of human embryonic stem cells. BMC Neurosci 2003;4:27.
29 Thomson JA, Itskovitz-Eldor J, Shapiro SS et al. Embryonic stem cell lines derived from human blastocysts. Science 1998;282:1145–1147.
30 Capi O, Gepstein L. Myocardial regeneration strategies using human embryonic stem cell-derived cardiomyocytes. J Control Release 2006; 116:211–218.
31 Kehat I, Kenyagin-Karsenti D, Snir M et al. Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. J Clin Invest 2001;108:407–414.
32 Xu C, Police S, Rao N et al. Characterization and enrichment of cardiomyocytes derived from human embryonic stem cells. Circ Res 2002;91:501–508.
33 Wilber A, Linehan JL, Tian X et al. Efficient and stable transgene expression in human embryonic stem cells using transposon-mediated gene transfer. Stem Cells 2007;25:2919–2927.
34 Chan KK, Wu SM, Nissom PM et al. Generation of high-level stable transgene expressing human embryonic stem cell lines using Chinese hamster elongation factor-1 α promoter system. Stem Cells Dev 2008; 17:825–836.
35 Chen CY, Schwartz RJ. Recruitment of the tinman homolog Nkx-2.5 by serum response factor activates cardiac α-actin gene transcription. Mol Cell Biol 1996;16:6372–6384.
36 Sepulveda JL, Belaguli N, Nigam V et al. GATA-4 and Nkx-2.5 coactivate Nkx-2 DNA binding targets: Role for regulating early cardiac gene expression. Mol Cell Biol 1998;18:3405–3415.
37 Christoforou N, Miller RA, Hill CM et al. Mouse ES cell-derived cardiac precursor cells are multipotent and facilitate identification of novel cardiac genes. J Clin Invest 2008;118:894–903.
38 Linhares VL, Almeida NA, Menezes DC et al. Transcriptional regulation of the murine Connexin40 promoter by cardiac factors Nkx2-5, GATA4 and Tbx5. Cardiovasc Res 2004;64:402–411.
39 Wang D, Chang PS, Wang Z et al. Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. Cell 2001;105:851–862.
40 Niwa H, Yamamura K, Miyazaki J. Efficient selection for high-expression transfectants with a novel eukaryotic vector. Gene 1991;108:193–199.
41 Sawicki JA, Morris RJ, Monks B et al. A composite CMV-IE enhancer/β-actin promoter is ubiquitously expressed in mouse cutaneous epithelium. Exp Cell Res 1998;244:367–369.
42 Xu ZL, Mizuguchi H, Ishii-Watabe A et al. Optimization of transcriptional regulatory elements for constructing plasmid vectors. Gene 2001;272:149–156.

See www.StemCells.com for supporting information available online.