HHS Public Access

Author manuscript

Gene Ther. Author manuscript; available in PMC 2018 December 21.

Published in final edited form as:

Gene Ther. 2018 August ; 25(5): 376–391. doi:10.1038/s41434-018-0021-z.

Stable and Reproducible Transgene Expression Independent of Proliferative or Differentiated State Using BAC TG-EMBED

Pankaj Chaturvedi, **Binhui Zhao**, **David L. Zimmerman***, and **Andrew S. Belmont**

Department of Cell and Developmental Biology, University of Illinois, Urbana, Illinois

Abstract

Reproducible and stable transgene expression is an important goal in both basic research and biotechnology, with each application demanding a range of transgene expression. Problems in achieving stable transgene expression include multi-copy transgene silencing, chromosome-position effects, and loss of expression during long-term culture, induced cell quiescence, and/or cell differentiation. Previously, we described the "BAC TG-EMBED" method for copy-number dependent, chromosome position-independent expression of embedded transgenes within a BAC containing ~170 kb of the mouse Dhfr locus. Here we demonstrate wider applicability of the method by identifying a BAC and promoter combination that drives reproducible, copy-number dependent, position-independent transgene expression even after induced quiescence and/or cell differentiation into multiple cell types. Using a GAPDH BAC containing ~200 kb of the human GAPDH gene locus and a 1.2 kb human UBC promoter, we achieved stable GFP-ZeoR reporter expression in mouse NIH 3T3 cells after low-serum induced cell cycle arrest or differentiation into adipocytes. More notably, GFP-ZeoR expression remained stable and copy-number dependent even after differentiation of mouse ESCs into several distinct lineages. These results highlight the potential use of BAC TG-EMBED as an expression platform for high-level but stable, long-term expression of transgene independent of cell proliferative or differentiated state.

INTRODUCTION

Transgene expression is an integral aspect of novel therapeutic regimes and production of mammalian antibodies, growth factors, cytokines, and DNA-based vaccines.[1] While many of these applications benefit from high-level transgene expression, in contrast other applications such as certain gene therapies may instead require low but stable levels of

Communicating Author: Andrew Belmont, asbel@illinois.edu, 217-244-2311, B107 CLSL, 601 S. Goodwin Ave, Urbana, IL 61801. Present Address: Biology Department, College of the Ozarks, Point Lookout, Missouri, United States of America

CONFLICT OF INTEREST: There are no competing financial interests in relation to this work.

Supplementary figure 1. GFP-ZeoR expression is maintained in clones with successive passages. (a) Reporter gene expression (y-axis, fluorescence levels normalized relative to day 0) for eleven NIH 3T3 cell clones carrying only reporter plasmid (UGG-Clone 01 to 11) over 28 days of continuous passaging (x-axis, days of culture) with zeocin selection (b) Reporter gene expression (y-axis, fluorescence levels normalized relative to day 0) for five mouse ES cell clones carrying GAPDH BAC (Clone 2-2, 4-4, 13 and 25) and only reporter plasmid (UGG-07) over 35 days of continuous passaging (x-axis, days of culture) with zeocin selection.

SUPPLEMENTARY INFORMATION:
Supplementary Movie and Tables are available online.

transgene expression. Optimal methods for transgene expression, therefore, should provide the ability to achieve both a reproducible and stable level of transgene expression. However, most commonly used approaches for transgene expression, and in particular transgene overexpression, result in unpredictable and unstable expression due to chromosome position effects and epigenetic gene silencing phenomenon.[2–4] Multi-copy plasmid-based transfection methods are particularly susceptible to these problems, which lead typically to copy-number independent expression levels highly variable between different cell clones as well as variegated expression within different cells from a single clone. This multi-copy transgene silencing makes plasmid-based systems unreliable for clinical or industrial research applications that need high-level, sustained expression of recombinant proteins in mammalian cells.

Commonly-used viral promoters for transgene expression such as CMV or SV40 have their highest activity in S-phase, and this activity decreases after induced cell quiescence.[5–9] Consequently, recombinant protein production from mammalian cells grown in bioreactors has been shown to be strongly proportional to cell growth rate.[10] However, specific productivity of monoclonal antibodies from hybridoma cultures is typically higher in growth-arrested cells.[10–12] Reduced transgene expression is also generally observed after induction of cell differentiation, which is often accompanied by prolonged or permanent cell-cycle arrest, as seen typically in *ex vivo* gene-therapy clinical applications. For example, *in vivo* extinction of transgene expression in transduced neural precursor cells has been observed in grafted tissues.[13] Similarly, unpredictable and unstable *in vivo* transgene expression in gene modified lymphocytes is a significant technical problem in cancer immunotherapy.[14] Silencing or “extinction” of reporter transgene expression during differentiation of embryonic stem cells has also been observed, producing considerable variation in transgene expression throughout the cell population.[15,16]

As reviewed elsewhere[4], attempts have been made to reduce these chromosome position effects on transgene expression by incorporating one or more *cis*-regulatory elements thought to help maintain a more transcriptionally-permissive chromatin environment into minigene constructs. These include scaffold-associated regions (SAR/MARs)[17,18], locus control regions (LCRs)[19], insulator or enhancer-blocking elements[20], or ubiquitous chromatin opening elements (UCOEs).[21,22] While inclusion of these regulatory sequences has been seen to improve expression of transgenic constructs[23], they may not confer both copy-number dependent and position-independent expression.[24]

In contrast, Bacterial Artificial Chromosomes (BACs) carrying large, ~100-200 kb mammalian genomic DNA insertions harbor most of the *cis*-regulatory sequences required for expression of the endogenous genes contained within these genomic inserts. Sustained copy-number dependent, position-independent expression of the genes contained within these mammalian genomic insertions has been seen in both cultured cells and transgenic animals after random chromosomal integration of these BACs.[25–28] Insertion of these large BAC transgenes appears to be sufficient to reconstitute large-scale chromatin organization similar to endogenous chromosome regions.[26,29] Moreover, BAC transgenes carrying the HBB or Hsp70 gene loci autonomously target to the nuclear periphery or nuclear speckles,

respectively, recapitulating the same nuclear compartment targeting of the endogenous HBB and Hsp70 gene loci.[29–31]

We therefore postulated that chromosome integration of these large mammalian genomic DNA inserts results in the reconstitution of a locally transcriptionally permissive chromatin environment largely independent of the chromosome integration site. We tested this hypothesis by inserting reporter minigenes at different locations within an ~170 kb mouse genomic DNA BAC insert containing the Dhfr gene locus, known to be active in proliferating cells. Using this "BAC TG-EMBED" method, we demonstrated stable reporter gene expression proportional to copy-number and independent of chromosome integration site over multiple cell passages for cell clones carrying several to hundreds of BAC copies.[26] Generating cell clones expressing transgenes at tens to hundred fold higher levels than a single-copy transgene using a single transfection step represents a potentially key technical improvement for transgene expression in diverse pharmacological and clinical applications where high-level expression is needed. Concurrently, similar results were obtained using a BAC containing the *Rosa*26 locus for expression of human IgG1-Fc fragment.[25]

However, the expression of Dhfr gene locus decreases or is silenced in non-proliferating cells.[32] Also, the viral promoters used for minigene expression in BAC TG-EMBED are prone to silencing over time via DNA methylation, particularly in certain cell types through a pathway involving DNA methylation.[16,33,34] Indeed, in the absence of selection we subsequently observed silencing of the CMV-driven reporter genes after long-term cell passaging.

Here we have modified the BAC TG-EMBED system specifically with the goal of testing the capability of this method to support long-term, stable expression in both dividing and terminally differentiated cells. We replaced the Dhfr BAC with a GAPDH BAC containing an ~200 kb genomic insert spanning the ubiquitously expressed GAPDH gene locus as the "scaffold" for inserting transgenes. We replaced the CMV promoter with an endogenous human UBC promoter that had earlier been shown to provide high and ubiquitous expression across many cell types.[35]

Using this new BAC and promoter combination, we now demonstrate the capability of BAC TG-EMBED to support high level, stable, copy-number dependent, chromosome position-independent reporter gene expression, which is maintained both through G0-arrest and cell differentiation. Specifically, we demonstrate sustained reporter gene expression after differentiation of fibroblasts into adipocytes and mouse ES cells into neuronal precursors and multiple cell types contained within embryoid bodies. Our results therefore demonstrate the potential of the BAC TG-EMBED for a large number of basic and therapeutic applications requiring high-level transgene expression stable through cell-cycle arrest in G0 and/or differentiation into a wide-range of cell types.

RESULTS

Construction of a new BAC TG-EMBED system for maintaining transgene expression after cell cycle arrest and/or cell differentiation

The Dhfr BAC used in our previous study can create a transcriptionally permissive chromatin environment in all proliferating cells, but it may not be an ideal BAC for the TG-EMBED method in terminally differentiated and/or quiescent cells where *Dhfr* gene is not expressed or is expressed at lower levels.[32] We therefore switched our BAC scaffold for the BAC TG-EMBED method to the RP11-369N23 BAC containing an ~200 kb human DNA genomic insert flanking the *GAPDH* (Glyceraldehyde 3-phosphate dehydrogenase) locus (GAPDH BAC). GAPDH is expressed widely in different tissue types and in both proliferating and non-proliferating cells.[36] Similarly, we chose the UBC promoter due to the ubiquitous expression of the UBC gene. To minimize silencing induced by unnatural DNA sequences found in foreign reporter and selection genes derived from non-mammalian species, we used a GFP-ZeoR fusion construct in which all CpGs had been removed to eliminate the possibility of DNA methylation. Here “GFP-ZeoR” refers to the construct expressing a fusion of GFP with the sh ble gene product conferring resistance to Zeocin.

We used BAC recombineering to insert a cassette (Figure 1a) containing the UBC-GFP-ZeoR minigene plus the GalK bacterial selectable marker into the GAPDH BAC. We used 74 bp homology arms to target this cassette into intron 1 of the *IFFO*1 gene, between nucleotide positions Chr12:6,661,887-6,661,888 (Figure 1b). GalK was used first with positive selection to insert this cassette into the BAC by homologous recombination using BAC recombineering, and then again with negative selection to isolate colonies in which the GalK gene flanked by FRT sites was excised using FLP recombinase.

The integrity of the recombineered BAC carrying the UBC-GFP-ZeoR reporter minigene was verified by restriction fingerprinting. BAC DNA was linearized through cutting the unique PI-SceI site in the vector backbone using the PI-SceI homing endonuclease prior to transfection of BAC DNA into mouse NIH 3T3 fibroblasts and zeocin selection of stable colonies (Figure 1c).

Individual stable clones were isolated and expanded for analysis of reporter gene expression and BAC transgene copy-number (Figure 1c). Flow cytometry revealed GFP-ZeoR fluorescence ranging 10-1000 fold higher than basal autofluorescence levels. Transgene copy-number was estimated by qPCR. Plotting mean fluorescence versus copy number for each clone revealed a linear relationship ($R^2 = 0.91$) between copy number and GFP-ZeoR reporter gene expression (Figure 2a). This linear relationship demonstrates the achievement of copy-number dependent, position-independent expression with this new BAC/reporter gene combination, as previously described with our original implementation of the BAC TG-EMBED method.[26] Transgene copy-numbers for these clones ranged from 1-77 with an average of 13 copies in two independent experiments.

The majority of NIH 3T3 cell clones (35/44) showed uniform GFP fluorescence levels among cells in the population, as demonstrated by a single peak spanning a roughly 10-fold range in fluorescence values in the flow cytometry histogram (Figure 2b, left). In contrast,

the remaining clones showed broad fluorescence distributions covering a 100-1000 fold range in fluorescence values (Figure 2b, right). Typically, these heterogeneous expression profiles were bimodal, with the lower fluorescence peak overlapping the autofluorescence peak seen in control cells (dashed lines, Figure 2b). This difference in expression patterns between different clones was confirmed by direct light microscopy visualization of GFP fluorescence within individual cells (Figure 2c).

However, both types of cell clones with either uniform (black circles, Figure 2a) or heterogeneous (red circles, Figure 2a) expression patterns showed a similar linear relationship between mean GFP fluorescence level and GFP-ZeoR reporter copy number. Given this tight correlation between expression level and mean BAC copy number, we attribute the heterogeneous expression profile to some type of genomic instability of BAC copy number in this subset of clones. A more detailed characterization of this genomic instability in mouse NIH 3T3 cells is in progress but beyond the scope of this current study. The GAPDH BAC produced ~80% uniformly expressing clones, as compared to other BAC scaffolds that produced ~40% uniformly expressing clones.[37] We focused our analysis on the uniformly expressing cell clones.

Long-term stability of BAC-embedded reporter gene expression in NIH 3T3 cells

Both epigenetic silencing and loss of amplified transgene loci can contribute to reductions in transgene expression during long-term culture.[20,38] Together they present a major challenge to maintaining a predictable and high yield of recombinant proteins. With our original BAC TG-EMBED system, we previously reported a 30-80% decline in reporter expression after long-term passaging in the absence of selection without any apparent loss of integrated BAC copies or large-scale chromatin reorganization.[26] We subsequently observed even larger declines of up to 90% for certain cell clones after long-term passaging for several months using our original BAC TG-EMBED system (unpublished data).

We tested the stability of reporter gene expression using our new BAC TG-EMBED system in five NIH 3T3 cell clones (3C9, 3F12, 4B2, 4C4 and 4E2). We passaged each cell clone carrying copies of the GFP-ZeoR reporter gene in the GAPDH BAC over a period of three months with or without zeocin selection, monitoring changes in GFP-ZeoR expression every 12 days by flow cytometry (Figure 3a). Cells were passaged every 3-4 days to maintain cells in a proliferative state and samples were analyzed by flow cytometry every 12 days for more than three months.

All 5 clones showed no significant decrease in reporter gene expression with or without zeocin selection (Figure 3a). One clone, 4E2, which was estimated as containing a single-copy insertion, exhibited a 2-fold or 3-fold increase in reporter gene expression with or without selection, respectively. For the remaining clones we observed small and variable increase in expression up to 1.5-fold over this 3-month period, possibly due to changes in growth conditions such as different serum lots during this long-term culture. We also maintained 11 NIH 3T3 cell clones carrying only reporter plasmid (UGG Cl-01 to 11) in proliferative state in zeocin selection over 4 weeks and did not observe any significant change in reporter expression (Supplementary Figure 1a).

Stability of BAC-embedded reporter gene expression in quiescent cells

We next were interested in examining the expression stability of the integrated, BAC-embedded reporter gene system in quiescent cells, a requirement common to most gene-therapy applications. Previous studies have reported reductions in transgene expression in G0-arrested cells, which has been attributed to the cell-cycle dependence of promoters such as CMV, SV40 and GADD153 (ref. [8, 40]) and/or alternatively to the remodeling / modification / condensation of chromatin, for instance in quiescent human T cells.[7]

To test the stability of expression in non-proliferating cells, we induced G0-arrest using a 96-hr exposure to low serum in five independent cell clones (2G6, 3C9, 3F12, 4B2 and 4C4) with integrated copies of the GAPDH BAC carrying the UBC-GFP-ZeoR reporter gene. We compared results using this new BAC TG-EMBED construct with results from cell clones containing either our original Dhfr BAC construct carrying a CMV-GFP reporter (clones GN-RZ-1, GN-RZ-2, GN-RZ-3 and GN-RZ-4)[26] or a Dhfr BAC construct carrying our new UBC-GFP-ZeoR reporter gene (clones F3-1, F3-15 and S1). A fraction of the cells used for reporter expression analysis was counterstained with propidium iodide (PI) to verify by flow cytometry the serum-starvation induced G0-arrest. The observed cell-cycle profiles showed <3.5% S-phase cells in these serum-starved cultures as compared to an estimated 17% in exponentially growing control cultures (Figure 3b).

Reporter gene expression in the G0-arrested GAPDH BAC clones varied from 62-104% of the level for that cell clone observed in exponentially growing cultures grown under normal serum conditions (Figure 3b). This compared to a slightly greater reduction in reporter gene expression in the Dhfr BAC clones carrying the original CMV-GFP reporter gene; expression levels ranged different clones from 38-58% of the expression level observed in exponentially growing cultures grown under normal serum conditions. This difference in the observed reduction in reporter gene expression with low serum between the GAPDH and Dhfr BAC scaffolds with the CMV-GFP reporter was small but statistically significant ($p < 0.05$). In contrast, clones containing Dhfr BAC scaffolds with the UBC-GFP-ZeoR reporter gene showed similar ranges in the reduction of reporter gene expression as seen for the GAPDH BAC reporter gene system, suggesting an effect of the reporter gene expression cassette rather than the BAC scaffold identity for the relative differences in response to low serum. The observed slight decrease in reporter gene expression could reflect a down-regulation of reporter gene transcription during G0-arrest but could also reflect a general reduction in translation or mRNA or protein stability associated with the G0-arrested state.

Stability of BAC-embedded reporter gene expression after differentiation of NIH 3T3 cells into adipocytes

A major problem with various gene therapy approaches is the silencing of transgenes accompanying cell differentiation. Chromosomal integration sites that support active transcription of transgenes in one cell type may not support continued transgene expression after differentiation to a different cell type.

Mouse NIH 3T3 cells are immortalized, non-tumorigenic cells with the potential to differentiate or transdifferentiate into cell types such as neural cells, osteocytes, myocytes,

adipocytes, or even insulin-producing beta cells either after genetic transformation or under defined media conditions.[39,40]

As a first test of the stability of BAC-embedded reporter gene expression after cell differentiation, we measured changes in GAPDH BAC-embedded GFP-ZeoR reporter gene activity after differentiation of NIH 3T3 fibroblasts into adipocytes (Figure 4a). Adipocyte differentiation was accomplished using a published protocol[40] which uses a PPARγ agonist (rosiglitazone) and an increased concentration of bovine growth serum to induce adipogenesis in NIH 3T3 cells.

Two days after reaching confluency, cultures of five GAPDH clones 2G6, 3C9, 3F12, 4B2 and 4C4 were then exposed to an adipogenic cocktail (0.5 mM methylisobutylxanthine, 1μM dexamethasone and 4.5 μM Rosiglitazone) for an additional 9 days. We tested the presence of adipocytes in the population after differentiation by Oil Red O staining of accumulated intracellular lipids. Most cells incubated in adipogenic medium stained positive for intracellular lipid vesicles, indicative of adipogenesis (Figure 4b). After differentiation into adipocytes, reporter gene expression in these 5 clones ranged from 31-76% (mean 48%) of the expression levels observed in control, exponentially growing cultures (Figure 4a).

BAC-embedded reporter expression in mouse embryonic stem cells

We next used mouse embryonic stem cells (mESCs) as a model system to explore more widely the ability of the BAC TG-EMBED system to sustain transgene expression after differentiation. mESCs are pluripotent and can be differentiated into representative cell types derived from all three germ layers. Using a transfection approach similar to that used in mouse NIH 3T3 cells, we isolated stable clones carrying GAPDH BAC transgenes in mESCs after zeocin selection (10 μg/ml) for two weeks.

A comparison of mean reporter gene fluorescence per copy number for 23 mESC clones (19 derived from C57BL/6N-PRX-B6N#1 and 4 derived from HM1 ES cells) revealed a similar linear relationship (R^2=0.8) (Figure 5a) as described previously for the same GAPDH BAC reporter gene combination in mouse NIH 3T3 cells (Figure 2a). Thus copy-number dependent, chromosome position-independent expression was also observed using the BAC TG-EMBED system in mESCs. In contrast to mouse NIH 3T3 cells, all mESC clones showed uniform expression of the GFP-ZeoR reporter gene, as measured by flow cytometry. This uniform expression pattern was visualized directly by microscopy of reporter gene expression within mESC colonies derived from two stable HM1 clones (Figure 5b).

The GFP-ZeoR fluorescence level per reporter gene copy, derived from the slope of the linear fit between mean fluorescence and copy number (Figure 5a), is ~30% lower than that seen for the same BAC construct in mouse NIH 3T3 cells (Figure 2a). This slightly lower expression may be related to the smaller cell size and/or faster cell cycle of mESCs relative to mouse NIH 3T3 cells. Similar to the results observed for NIH 3T3 clones, we did not observe any change in expression of ES cell clones (GAPDH 2-2, 4-4, 13 and 25, and UGG-07) with continuous passage over 35 days (Supplementary Figure 1b).

Stability of BAC-embedded reporter gene expression in mESCs differentiated into neural progenitor cells

Having established these mESC clones, we next evaluated the dependence of BAC-embedded reporter gene expression after differentiation using two well-characterized mESC differentiation protocols: 1) Neural induction of monolayer cells using all *trans*-retinoic acid (RA) treatment, 2) Non-directed multi-lineage differentiation by embryoid body (EB) formation.

We used C57BL/6N-PRX-B6N#1 mESC Clone 13 showing a uniform expression distribution for directed differentiation into neuronal progenitor cells using a *trans*-retinoic acid (RA) treatment protocol described previously.[41] Briefly, mESCs were induced to form cellular aggregates (CA) and treated with 5 μM all *trans*-retinoic acid for 4 days. CAs maintained in suspension for another 4 days were then disaggregated and plated on gelatinized dishes or coverslips. After two weeks of differentiation cells were analyzed for expression of β-III tubulin, a differentiation marker for neural cells; β-III tubulin expression is associated with differentiation and decreased cell proliferation in neurons and neuronal/glial precursors.[42] In contrast to undifferentiated ES cells, these RA exposed cells showed strong expression of β -III tubulin (Figure 5c). In four independent differentiation experiments, the efficiency of neural differentiation varied from 60-70%.

All cells expressing β-III tubulin were also positive for the GFP-ZeoR reporter (Figure 5c). Flow cytometry showed GFP-ZeoR fluorescence levels between ~80-120% of the corresponding values in the parent, undifferentiated mESCs (Figure 5d). Furthermore, we also compared flow cytometry profiles of mESC GAPDH BAC clones (2-2, 4-4 and 13) and reporter plasmid clones (01, 05 and 07) after neural differentiation (Figure 5e).

No clone carrying the GAPDH BAC transgenes showed decreases in reporter gene expression after neuronal differentiation. Changes in fluorescence varied from nearly no change in reporter gene expression (Figure 5e, Clone GAPDH-13) to a 2-3-fold increase in mean reporter gene expression (Figure 5e, Clones GAPDH 2-2, 4-4) after neuronal differentiation. In striking contrast, in two cell clones (UGG-01 and 05) carrying the plasmid reporter genes, an ~10-100-fold reduction of reporter gene expression was observed in nearly the entire cell population. In the third clone (UGG-07), a strong variegated reporter gene expression was observed, with a substantial fraction of the cell population also reduced ~10-fold after neuronal differentiation (Figure 5e, right 3 panels).

BAC TG-EMBED reporter gene expression is maintained within multiple types of spontaneously differentiated cells within embryoid bodies

In vitro differentiation of ES cells can be induced by aggregating cells into embryoid bodies (EBs) and maintaining them in suspension on a non-adherent surface to mimic early stages of embryonic lineage specification and cellular differentiation. Many different cell types from all three germ layers are present within EBs.

Six individual ES cell clones (GAPDH clones 2-2, 4-4, and 13, and UGG reporter clones 01, 05 and 07) were induced to differentiate spontaneously as EBs. After two weeks of non-

directed, multi-lineage EB differentiation, clones were analyzed for GFP-ZeoR reporter expression by flow cytometry.

All 3 ES cell clones carrying the GAPDH BAC transgenes showed cells at day 14 of EB differentiation with GFP-ZeoR fluorescence levels comparable to levels in undifferentiated ES cells (Figure 6a). In two clones the fluorescence distribution remained either remain nearly the same (GAPDH-13) or increased in mean fluorescence ~70% (GAPDH-4-4) after differentiation. In the third clone (GAPDH-2-2) the mean fluorescence also increased ~70% but the distribution became broader after differentiation.

In striking contrast, again all three clones carrying the plasmid version of the reporter gene transgene showed a pronounced reduction of reporter gene expression in near 100% of the cell populations, with mean fluorescence levels reduced from 12-100-fold between the different cell clones (Figure 6a).

Two of the differentiated clones (GAPDH-2-2 and 4-4) carrying the GAPDH BAC transgenes were further stained with different cell lineage specific markers: β-III tubulin (neurons and neuronal/glial precursors) and nestin (neural stem/progenitor cells) (both neuroectoderm), cardiac troponin T (cardiomyocytes) and myosin heavy chain (skeletal and cardiac muscle) (both mesoderm), and SOX17 (definitive endoderm).

All cells staining positive for nestin showed GFP fluorescence at very similar levels of intensity in both cell clones (Figure 6b-c). Similarly, cells positive for β-III tubulin, showed uniform GFP fluorescence levels. Furthermore, we observed that by day 10 of differentiation more than 75% of adherent EBs showed spontaneous contractile activity. These EBs had cell clusters expressing the cardiomyocyte marker cardiac troponin T (cTroponin) and showed spontaneous “beating” contractions. These cardiomyocyte cell clusters showed higher GFP-ZeoR reporter gene expression compared to undifferentiated cells and the other differentiated cells in embryoid bodies (Figure 6b-c, Supplementary Movie 1). GFP-ZeoR fluorescence levels in cells staining positive for skeletal myosin heavy chain (skMHC), a marker for skeletal muscle cells, were also comparable to levels in undifferentiated cells and other differentiated cells within the EBs (Figure 6b-c).

The only cell marker examined which showed cells with variable GFP-ZeoR expression was SOX17. SOX17 is a well-characterized endoderm marker and is required for definitive gut endoderm formation in mice and other species.[43] The differentiated EBs immunostained for SOX17 showed patches of SOX17-positive cells that contained both GFP positive and negative cells (Figure 6b-c). We speculate that the fully differentiated primitive endoderm cell subtype may be negative for reporter gene expression using this GAPDH BAC / UBC promoter TG-EMBED system while the GFP-positive cells may represent partially differentiated cells on the pathway to terminally differentiated, primitive endodermal cells.

DISCUSSION

In summary, we have demonstrated the ability of the BAC TG-EMBED method to maintain long-term, copy-number dependent and chromosome position-independent expression of transgenes after induction of cell quiescence and/or cell differentiation. Using a BAC

containing the ubiquitously expressed GAPDH locus and the human UBC promoter to drive expression of a CpG-free GFP-ZeoR reporter gene, we demonstrated copy-number dependent, chromosome position-independent expression in mouse NIH 3T3 fibroblasts with reporter gene fluorescence levels dropping no more than 40% after induction of cell cycle arrest through low serum. Differentiation of these same NIH 3T3 cell clones into adipocytes resulted in no more than a 2-3-fold drop in fluorescence among multiple clones. Similarly, reporter gene fluorescence remained stable using this same BAC TG-EMBED system in mouse ESCs differentiated into a range of different cell types, but not for the plasmid reporter transgene where we observed ~10-100-fold decreases in expression. Differentiation into neuronal precursor and neuronal cells resulted in a less than 30% decrease in reporter gene fluorescence. Reporter gene fluorescence was also maintained in the vast majority of cells within embryoid bodies differentiated into multiple cell types, including neuronal and both cardiac and skeletal muscle cell types. Of the cell types assayed by several lineage markers, only primitive endoderm-type cells showed reporter gene variegation.

This maintenance of transgene expression after induction of cell quiescence and/or cell differentiation represents a marked improvement relative to that typically observed using existing transgene expression technology. For example, lentivirus currently is one of the preferred vector delivery systems for high-level, sustained transgene expression. Even using lentivirus, only a fraction of integration sites are permissive for transcription; other sites show either silencing or variegation.[16] Moreover, integration sites permissive for expression of the lentivirus transgene may not be permissive after differentiation. For example, lentivirus transduction of GFP driven by the strong EF1α promoter into mESCs, EBs, NPCs, and fully differentiated neuronal cells produced similar levels of GFP fluorescence that persisted over extended culture time in all cell types. However, if either undifferentiated mESCs or neuronal precursor cells were transduced with this GFP lentivirus and then differentiated into neuronal cells, transgene silencing led to a roughly 10-fold reduction in transgene mRNA levels.[44] Similarly, the majority of GFP-positive cell clones carrying multiple copies of a lentivirus GFP reporter showed only 3-14% GFP positive cells after hematopoietic differentiation (with no information available concerning the actual GFP fluorescence level in the remaining GFP+ cells).[45]

In contrast, using BAC TG-EMBED 100% of colonies showed reporter gene expression with copy-number dependent expression that did not show significant change during cell quiescence and most differentiated states. Even in the case of differentiation of NIH 3T3 fibroblasts to adipocytes, where we saw an ~3-fold drop in reporter gene fluorescence per cell, all colonies retained fluorescence after differentiation.

Currently, the newest and most promising method to ensure stable expression of transgenes is by using genome editing tools to insert transgenes in transcriptionally permissive loci or “hot spots” such as mouse *ROSA26*, *Hprt1*, *H11* or human adeno-associated virus integration site 1 (*AAVS1*). These “safe haven” loci are known to have 10-100 times higher reporter expression.[46–49] Yet even at these sites, expression of transgenes may be too low for experimental needs. For example, insertion of a single Rosa26 BAC in which a GFP reporter was placed in the first exon of transcript 1 of this locus produced higher fluorescence than

was reported by insertion of a GFP reporter into the endogenous Rosa26 locus, possibly due to loss of adjacent repressive sequences[50], and other anecdotal reports have suggested that targeting a reporter gene such that it will be under control of an endogenous promoter may produce too low expression for visualization of reporter gene activity.[51] Additionally, silencing of both the endogenous (EF1α) and viral (CMV) promoters at the *AAVS1* locus in human iPSCs has been reported.[52]

In comparison, our new BAC TG-EMBED method has several advantages. It can be applied to cell types in which genome editing has low efficiency. Promoters driving transgene expression using the BAC TG-EMBED method can be chosen based on their expression level or ability to be induced. This is different from approaches in which transgenes are placed under control of promoters of endogenous genes lying within the safe havens. The range of transgene expression can be further varied up to ~100-fold based on the number of integrated BAC copies; this is particularly valuable when high levels of expression are desired. Finally, we anticipate, as demonstrated previously[26], that multiple transgenes will be able to be expressed simultaneously at predictable relative levels of expressed by embedding them within the same BAC scaffold.

Another difference from our original BAC TG-EMBED work is that we deliberately used a GFP-ZeoR reporter gene body containing no CpG dinucleotides together with the human UBC promoter sequence. DNA methylation at CpGs in promoter and gene body sequences contributes to one form of transgene silencing. CpG-free transgenes have been exploited in therapeutic applications to provide more predictable expression; reduction or elimination of CpGs from plasmid DNA improve stability of transgene expression over periods up to 60 days in mouse lungs.[53] However, although CpG-free sequences can ameliorate silencing caused by DNA methylation, non-mammalian minigene sequences can still act as nucleation centers for heterochromatin formation and its spread in *cis.*[54] This heterochromatin formation initiates at plasmid backbone sequences and over a period of week or more completely silences the transgene unless the spread is countered by insulator sequences.[55,56] We used a CpG-free reporter gene body to eliminate any contribution to transgene repression coming from such DNA methylation such that we could examine the inherent ability of the BAC TG-EMBED method to shield sequences from other types of heterochromatin formation associated with either chromosome position effects or from unnatural features of minigene sequence composition.

Our working hypothesis is that the various *cis*-regulatory regions distributed over the GAPDH BAC scaffold are sufficient to reconstitute a normal large-scale chromatin environment permissive for transgene expression independent of the integration site. Transcription at one site may lead to a more permissive site in flanking regions, as inferred from a transcription “ripple” effect in which increased expression of some flanking genes up to 200 kb away is seen after activation of immediate early genes.[57] A genome-wide study assaying activity of randomly integrated reporter in mammalian cells also suggested that expression of transgenes is higher when it is located within 200 kb of highly active genes.[58] In previous work, we have shown that BAC transgenes reconstitute a large-scale chromatin compaction comparable to endogenous euchromatic chromosome regions, as visualized by both light and electron microscopy.[29,30]

In conclusion, we have established that BAC TG-EMBED maintains stable and predictable copy-number dependent, chromosome position-independent minigene expression both after induction of cell quiescence or large changes in differentiated cell state. This ability combined with the ability to tune transgene expression over two orders of magnitude by varying the copy number of integrated BAC copies, should make the BAC TG-EMBED an attractive method for transgene expression for certain biopharmaceutical and therapeutic applications. Our-long term goal is to be able to engineer entire synthetic gene networks into human cells for improved *ex vivo* gene therapy and tissue engineering applications. Based on our previous work showing the feasibility of multi-transgene expression on a single BAC scaffold[26], we anticipate that the BAC TG-EMBED method will also prove useful for stable, reproducible expression of multiple transgenes during induction of cell quiescence and/or cell differentiation.

MATERIALS AND METHODS

Vector construction

Different preexisting plasmids designed for other experiments were repurposed for construction of pUGG. The CpGfree GFP-ZeoR-SV40 polyA fragment was originally derived from p[Select-zeo-GFPSh] (nucleotide position- 724-2056, InvivoGen). p[Select-zeo-GFPSh] contains an expression cassette (GFPSh) containing a red-shifted variant of GFP fused to the *Sh ble* gene conferring resistance to Zeocin. We PCR amplified the fragment from p[CpGfree-vitroNeo-ZGFP] plasmid (a gift from Peter Jones, University of Nevada Medical School), where ZGFP is a renaming of GFPSh, with the following flanking sequences added to the fragment (5’ end: 5’-ACCCTGCAGCCCTAGACAATTGTACTAACCTTCTTCTCTTTCCTCTCCTGACAGGTTGGTGTACAGTAGCTTC-3’; 3’ end: 5’-GGAGCCCCACTGTGTTCATCTTACAGATGGAAATACTGACATTCAGAGGAGTTAGTTAACTTGCCTAGGTGATTCAGCTAATAAGTGCAAGAAAGATTTCAATCCAAGGTGATTTGATTCTGAAGCCTGTGCT-3’; PCR primer binding regions are underlined). The 1.6 kb PCR amplified fragment was cloned into XhoI, EcoRV sites of pZeo[26] creating an intermediate plasmid p[ZGFP-Zeo]. The original SV40 promoter-Zeocin-polyA cassette (1.0 kb) of pZeo was then deleted from pZGFP-Zeo using XhoI and BamHI, followed by religation of blunted ends to create pZGFP. A 2.4 kb CpGfree SV40-Neo cassette, excised from p[CpGfree-vitroNeo] (InvivoGen), was then introduced into pZGFP at EcoRI/EcoRV sites to get intermediate plasmid p[SV40-Neo-ZGFP]. Human Ubiquitin-C promoter (hUBC) (1.3 kb) was PCR amplified from plasmid pFUGW[35] and cloned into the EcoRV site in pSV40-Neo-ZGFP plasmid to produce pSNUG (pSV40-Neo-Ubiquitin-GFP Zeocin).

To construct a donor vector for BAC recombineering a 2.85 kb fragment carrying SV40 promoter-CpGless Neo was first eliminated from plasmid pSNUG by double digestion with ClaI/BglII restriction enzymes. The remainder of the plasmid was ligated with Fragment I (Supplementary Table 1) to recircularize the plasmid (renamed pUG). For targeted recombination using a galK positive/counter selection, a galK cassette flanked by FRT sites was introduced into the plasmid between the HindIII and SpeI sites. The vector pUG was digested with SphI/XhoI. Fragments II and III (Supplementary Table 1) bearing FRT sites,

and either XhoI/HindIII or SpeI/SphI overhangs were first ligated to the two ends of the linearized pUG. The 1.25 kb GalK DNA fragment excised from plasmid pGalK[59] by HindIII/SpeI double digestion was then ligated with the linear pUG plasmid with FRT sites at each terminus. The resulting construct pUGG carries the Galk cassette flanked by FRT sites for recombination. All the restriction enzymes used have been procured from New England Biolabs (NEB), UK.

Constructing BACs containing the UBC-GFP-ZeoR reporter

BAC RP11-369N23 carrying the human GAPDH genomic locus (GAPDH BAC) was obtained from Life technologies. BAC CITB-057L22 (Dhfr BAC) was a gift from Edith Heard (Curie Institute, Paris, France) and contains mouse chr13: 92,992,156-93,161,185 (NCBI37/mm9). Genomic coordinates and sequences described in this study are based on the hg19 and mm9 genome assemblies as reported in the UCSC Genome Browser (http://genome.ucsc.edu/).[60] λ red–mediated BAC recombineering using a *galK*-based dual-selection scheme was used to introduce the UBC-GFP-ZeoR reporter cassette between the nucleotide positions Chr12: 6,661,887-6,661,888 on the GAPDH BAC and Chr13:93,099,101-93,099,102 on Dhfr BAC.[59] BAC DNA was transformed into *Escherichia coli* strain SW105 in which the λ red recombination machinery is induced by raising the temperature from 32 to 42°C.[59,61] DNA fragments with homology arms were prepared by PCR using long primers (Supplementary Table 2) with 74-bp homology sequences against the BAC target integration site plus 16-bp sequences (forward, 5'- acagcagagatccagt-3'; reverse, 5'-tgttggctagtgcgt-3') to amplify the UBC-GFP-ZeoR-Galk cassette from plasmid pUGG. Recombinants were selected for *galK* insertion at 32°C on minimal medium in which D-galactose (Sigma) was supplied as the only carbon source. Recombinant colonies were PCR screened using BAC specific primers flanking the target regions (Supplementary Table 2). *galK* was removed from recombinant clones by inducing FLP-recombinase in actively growing SW105 cells using 0.1% (w/v) L-arabinose (Sigma). Negative selection used minimal medium containing 2-deoxy-galactose (Sigma), and deletion of *galK* in recombinants was verified using BAC specific primers (Supplementary Table 2). Integrity of BAC constructs was verified by restriction enzyme fingerprinting, by comparing observed band patterns on agarose gel to the digestion patterns predicted using Gene Construction Kit (Textco BioSoftware) or ApE software (http://biologylabs.utah.edu/jorgensen/wayned/ape/). The final BAC constructs with the UBC-GFP-ZeoR cassette inserted and galK removed (GAPDH+UG) or (Dhfr+UG) were prepared for transfection into mammalian cells using the Qiagen Large Construct Kit.

Cell culture and establishment of NIH 3T3 cell clones with integrated BAC or plasmid DNA

NIH 3T3 cells (CRL-1658, ATCC) were routinely grown in Dulbecco's modified Eagle medium (Life technologies) supplemented with 10% Bovine Growth Serum (BGS) (HyClone) and 1X Antibiotic-Antimycotic (Life Technologies). Purified GAPDH BAC DNA was linearized using the PI-SceI homing endonuclease and transfected into NIH 3T3 cells using Lipofectamine 2000 (Life technologies) according to the manufacturer's directions. The reporter plasmid pUGG was linearized with NdeI and gel purified prior to transfection. Mixed clonal populations of stable transformants were obtained after 2 weeks

of 75 μg/ml zeocin (Life Technologies) selection; individual cell clones were obtained by serial dilution or retrieval of colonies using filter discs.[62]

Flow cytometry

GFP-ZeoR reporter gene expression levels were analyzed using a FACS Canto Flow Cytometry Analyzer (BD Biosciences) using a 488nm laser to detect GFP and/or Propidium Iodide (PI) through 530/30 or 625/635 nm LP emission filters, respectively. Rainbow fluorescent beads RFP-30-5A (Spherotech Inc.) were used as an intensity standard to normalize GFP fluorescence; the flow cytometer PMT voltage was adjusted to produce the same output intensity for these beads prior to each use of the instrument. Mean cell fluorescence intensity (MFI, in arbitrary units) of GFP normalized to fluorescent bead intensity was used as a measure of transgene expression. To ensure uniform normalization for all samples, fluorescent beads from the same batch were used for all measurements. Untransfected NIH 3T3 and mouse ES cells were used to establish background cell auto-fluorescence levels. Linear fitting of mean GFP expression level versus transgene copy-number for each group of cell clones was performed using Microsoft Excel and fixing the y-intercept "*a*" to the fluorescence background level of non-transfected cells. The correlation coefficient R^2 when the y-intercept is fixed is defined as:

$$R^2 = bb', \text{ where } b = (\sum x_i y_i - a\sum x_i)/\sum x_i^2 \text{ and } b' = (\sum x_i y_i - a\sum x_i)/\sum (y_i - a)^2.$$

Transgene copy-number measurement

Genomic DNA from individual cell clones was isolated using phenol-chlorofom-isoamyl alcohol extraction. DNA concentration was measured by UV absorbance using a Nanodrop instrument (Thermo Scientific). BAC or plasmid transgene copy-number within individual cell clones was determined by real-time quantitative PCR on a PCR instrument (StepOnePlus, Applied Biosystems) using iTaq universal SYBR green supermix (Bio-Rad). Each assay was done in a final volume of 20 μL, consisting of 10 μL 2x SYBR Green Master Mix (Bio-Rad), 10 ng genomic DNA template, 100 nM each primer, and nuclease free water. Samples were run in triplicate in 96 well plates with the following cycling parameters: 95 °C for 10 min, then 40 cycles of 95 °C for 15 s, 58 °C for 30 s and 72 °C for 30 s. The amplification phase was followed by a dissociation step at 95 °C for 15 s, 60 °C for 1 min and 95 °C for 15 s to check the purity of the amplification products. The mouse serum/glucocorticoid regulated kinase 1 (*Sgk1*) gene was used as an internal control to estimate number of transgene copies making an assumption that the number of copies of *Sgk1* in tetraploid NIH 3T3 cells is 4 and in diploid mouse ES cells the copy number is 2. The efficiency of primers used for amplification ranged from 95% to 105%. Primers used to measure BAC copy-number and reference genes are listed in Supplementary Table 2.

NIH 3T3 cell cycle arrest induced by serum starvation

Low serum was used to induce G0-arrest of NIH 3T3 cell clones carrying GAPDH BAC transgenes.[63] NIH 3T3 cells at 60-70% confluency were maintained in DMEM supplemented with 0.2% BGS (HyClone) for 96 hrs. Cell cycle arrest was assayed by flow cytometry. Cells were detached by trypsinization, pelleted by centrifugation, resuspended in

1XPBS, fixed in ice-cold 70% (v/v) ethanol, and then stored in this 70% ethanol at 4°C. Prior to DNA content analysis, cells were washed in PBS, followed by staining with 2 μg/ml propidium iodide (Sigma) in 1XPBS supplemented with 1 μg/ml DNase-free RNase A (Sigma) at 37°C for 1hour. PI fluorescence staining of DNA was measured on a FACS Canto Flow Cytometry Analyzer (BD Biosciences). Ten thousand events were acquired for each sample. The percentage of diploid cells in G1, S, and G2 was estimated by FACSdiva software (BD Biosciences). FACS analysis of PI stained, serum starved cells revealed >90% of cells with G1 DNA content; the estimated fraction of S phase cells was <4%, in contrast to asynchronously growing control cells which had >15% cells in S-phase.

Adipocyte differentiation of NIH 3T3 cell clones

To induce adipocyte differentiation, NIH 3T3 cells which had been confluent for 48 hrs were incubated for 9 days in differentiation medium (DMEM plus 20% BGS supplemented with MDR (0.5 mM methylisobutylxanthine, 1μM dexamethasone and 4.5 μM Rosiglitazone) as described elsewhere.[40] To assay the percentage of cells differentiated into adipocytes, differentiated and control cells were stained using the Oil Red O method. Briefly, cells in 6-well plates were washed twice with PBS, fixed in 3.7% formaldehyde for 30 mins, and washed with double deionized H_2O (ddH_2O). After 5 min incubation with 60% isopropanol in ddH_2O (v/v), cells were stained for 5 mins in 1 ml of freshly prepared 0.3% Oil Red O (0.3 gm Oil Red O dissolved in 60 ml isopropanol, filtered, and then brought to 100 ml with ddH_2O), then washed with ddH_2O until the water rinse was clear.

Mouse embryonic stem cell maintenance, establishment of cell clones, and differentiation

Undifferentiated mouse ES cells C57BL/6N-PRX-B6N#1 (Jackson Labs) and HM1 (provided by A. Smith, University of Cambridge, Cambridge, England, UK) were maintained on a feeder layer of mitomycin C–treated mouse embryonic fibroblasts (MEFs) or in flasks coated with 0.1% gelatin (Sigma) in ES cell medium DMEM (Gibco) supplemented with 15% fetal bovine serum (FBS; HyClone), 1% L-glutamine 200mM (Gibco), 1X Antibiotic-Antimycotic (Life technologies), 1% MEM non-essential amino acids (Gibco), 0.1mM 2-Mercaptoethanol (Sigma) and 1000 Units/ml ESGRO-LIF (Millipore). MEFs were obtained from the American Type Culture Collection and maintained in DMEM supplemented with 10% FBS (Hyclone).

Purified GAPDH BAC DNA was linearized using the PI-SceI homing endonuclease and transfected into C57BL/6N-PRX-B6N#1 (Jackson Labs) or HM1 mouse ES cells using Lipofectamine 2000 (Life technologies) according to the manufacturer's directions. The reporter plasmid pUGG was linearized with NdeI and gel purified prior to transfection. Mixed clonal populations of stable transformants were obtained after 2 weeks of 10 μg/ml zeocin (Life Technologies) selection; individual cell clones were obtained by retrieval of isolated colonies using wide-bore pipette tips. The stable clones were then expanded on 0.1% gelatin coated dishes, and screened for reporter expression.

To derive neuronal cells, murine ES cells were allowed to aggregate on non-adherent dishes at a density of 4×10^6 ES cells/ml in differentiation medium consisting of ES cell medium lacking LIF.[41] Cellular aggregates in differentiation media were supplemented with 5 μM

all-*trans*-Retinoic acid (Sigma) from day 4 onwards and maintained in suspension for another 4 days. After these 8 days of differentiation, aggregates were dissociated, plated on 0.1% gelatin-coated dishes in Neurobasal medium (Gibco) supplemented with N2-Supplement (Life Technologies) and cultured for another 10 days.

For spontaneous differentiation into embryoid bodies (EBs), ESCs cultured on gelatin-coated dishes were trypsinized with 0.25% trypsin-EDTA (Gibco), washed twice in 1X Phosphate Buffer Saline (PBS) and resuspended in differentiation medium. Cells were incubated (6×10^3 cells per 30 μl droplet) using the hanging drop method to form EBs as described elsewhere.[64] EBs were maintained in differentiation medium in suspension culture for 4 days (2 days as hanging drops and another 2 days in bacteriological grade petri dishes), followed by adherence on 0.1% gelatin-coated tissue culture-grade petri dishes or 4-chamber slides (Nunc Lab-Tek) for another 6-10 days. Differentiation media was replenished after every 2 days in culture.

Immunofluorescence

Mouse monoclonal antibodies to cardiotroponin T (1:500, Catalog #ab10214; Abcam), nestin (1:1000, Catalog #ab6142; Abcam), SOX17 (1:200, Catalog #ab192453; Abcam), skeletal myosin heavy chain (1:40, Catalog #MF20; Developmental Studies Hybridoma Bank) and rabbit polyclonal antibody to β-III tubulin (1:1000, ab18207; Abcam) were used. For immunofluorescence assays, EBs on day 5 were transferred onto 0.1% gelatin-coated glass coverslips placed in 6-well tissue culture plates. After 10 days of culture, EBs were fixed for 15 min with 4% paraformaldehyde (Sigma) (w/v) in phosphate-buffered saline (PBS) followed by permeabilization in PBS with 0.5% Triton X-100 (Sigma) (v/v) for 1 hr at room temperature. Cells were incubated with 0.5% normal goat serum in PBS for 3-4 hrs at room temperature. The EBs were then incubated with primary antibody diluted in 1XPBS overnight at 4°C, washed 3x 5 mins in 1XPBS, and then incubated in anti-mouse IgG or anti-rabbit IgG Texas Red-conjugated secondary antibody (1:500, Catalog #115-075-006 or 111-075-006; Jackson Immunoresearch) for 1 hr at room temperature in 1XPBS. Coverslips were washed 3x in 1XPBS and mounted in antifade medium containing 0.3 mg/ml DAPI (Sigma)/10% w/v Mowiol 4–88 (EMD)/1% w/v DABCO (Sigma)/25% glycerol/0.1 M Tris, pH 8.5. Fluorescent images of the stained EBs or cells were acquired with a Deltavision deconvolution microscope system (Applied Precision) with a 60X/1.4 NA lens or LSM700 confocal microscope (Zeiss) with a NeoFluar 20X/0.5 NA DIC or 40X/1.30 NA Oil DIC lens for data collection. Deconvolution of images acquired using Deltavision microscope was done by an enhanced ratio, iterative constrained algorithm[65], using the Applied Precision SoftWorx software. Images were analyzed with ImageJ software and assembled using Adobe Photoshop or Adobe Illustrator.

Statistical analyses

All statistical analyses were carried out using Excel (Microsoft) and GraphPad Prism (GraphPad Software). Unless otherwise stated, data correspond to mean±SEM (standard error of the mean). P-values were calculated using two-tailed unequal variance *t* tests.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by grant R01 GM098319 from the National Institute of General Medical Sciences (ASB). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of General Medical Sciences or the National Institutes of Health. We thank Peter Jones (University of Nevada Medical School) for providing pCpGvitro-neo-ZGFP and T. Brennan (University of California, San Francisco, San Francisco, CA) for FUGW plasmids. We also thank A. Smith (University of Cambridge, Cambridge, England, UK) for HM1 ES cells. We are also grateful to Edith Heard (Curie Institute) for providing Dhfr BAC (clone 057L22 from CITB mouse library).

Support: NIH funded

References

1. Walsh G. Biopharmaceutical benchmarks. Nat Biotechnol. 2006; 24:769–776. [PubMed: 16841057]
2. Karpen GH. Position-effect variegation and the new biology of heterochromatin. Curr Opin Genet Dev. 1994; 4:281–291. [PubMed: 8032206]
3. Dorer DR, Henikoff S. Transgene repeat arrays interact with distant heterochromatin and cause silencing in cis and trans. Genetics. 1997; 147:1181–90. [PubMed: 9383061]
4. Kwaks THJ, Otte AP. Employing epigenetics to augment the expression of therapeutic proteins in mammalian cells. Trends Biotechnol. 2006; 24:137–42. [PubMed: 16460822]
5. Zopf CJ, Quinn K, Zeidman J, Maheshri N. Cell-cycle dependence of transcription dominates noise in gene expression. PLoS Comput Biol. 2013; 9:e1003161. [PubMed: 23935476]
6. Akita H, Ito R, Kamiya H, Kogure K, Harashima H. Cell cycle dependent transcription, a determinant factor of heterogeneity in cationic lipid-mediated transgene expression. J Gene Med. 2007; 9:197–207. [PubMed: 17351985]
7. Cooper LJN, Topp MS, Pinzon C, Plavec I, Jensen MC, Riddell SR, et al. Enhanced Transgene Expression in Quiescent and Activated Human CD8+ T Cells. Hum Gene Ther. 2004; 15:648–658. [PubMed: 15242525]
8. Brightwell G, Poirier V, Cole E, Ivins S, Brown KW. Serum-dependent and cell cycle-dependent expression from a cytomegalovirus-based mammalian expression vector. Gene. 1997; 194:115–123. [PubMed: 9266680]
9. Dutton RL, Scharer J, Moo-Young M. Cell cycle phase dependent productivity of a recombinant Chinese hamster ovary cell line. Cytotechnology. 2006; 52:55–69. [PubMed: 19002865]
10. Lee FWF, Elias CB, Todd P, Kompala DS. Engineering Chinese hamster ovary (CHO) cells to achieve an inverse growth–associated production of a foreign protein, β-galactosidase. Cytotechnology. 1998; 28:73–80. [PubMed: 19003409]
11. Suzuki E, Ollis DF. Cell Cycle Model for Antibody Production Kinetics. Biotechnol Bioeng. 1989; 34:1398–1402. [PubMed: 18588084]
12. Fussenegger M, Schlatter S, Dätwyler D, Mazur X, Bailey JE. Controlled proliferation by multigene metabolic engineering enhances the productivity of Chinese hamster ovary cells. Nat Biotechnol. 1998; 16:468–72. [PubMed: 9592397]
13. Rosenqvist N, Hård Af Segerstad C, Samuelsson C, Johansen J, Lundberg C. Activation of silenced transgene expression in neural precursor cell lines by inhibitors of histone deacetylation. J Gene Med. 2002; 4:248–257. [PubMed: 12112642]
14. Burns WR, Zheng Z, Rosenberg SA, Morgan RA. Lack of specific γ-retroviral vector long terminal repeat promoter silencing in patients receiving genetically engineered lymphocytes and activation upon lymphocyte restimulation. Blood. 2009; 114:2888–2899. [PubMed: 19589923]
15. Laker C, Meyer J, Schopen A, Friel J, Heberlein C, Ostertag W, et al. Host cis-mediated extinction of a retrovirus permissive for expression in embryonal stem cells during differentiation. J Virol. 1998; 72:339–348. [PubMed: 9420232]

16. Yao S, Sukonnik T, Kean T, Bharadwaj RR, Pasceri P, Ellis J. Retrovirus silencing, variegation, extinction, and memory are controlled by a dynamic interplay of multiple epigenetic modifications. Mol Ther. 2004; 10:27–36. [PubMed: 15233939]

17. Phi-Van L, von Kries JP, Ostertag W, Strätling WH. The chicken lysozyme 5' matrix attachment region increases transcription from a heterologous promoter in heterologous cells and dampens position effects on the expression of transfected genes. Mol Cell Biol. 1990; 10:2302–7. [PubMed: 2325653]

18. Kim JM, Kim JS, Park DH, Kang HS, Yoon J, Baek K, et al. Improved recombinant gene expression in CHO cells using matrix attachment regions. J Biotechnol. 2004; 107:95–105. [PubMed: 14711493]

19. Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell. 1987; 51:975–985. [PubMed: 3690667]

20. Pikaart MJ, Recillas-Targa F, Felsenfeld G. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. Genes Dev. 1998; 12:2852–2862. [PubMed: 9744862]

21. Antoniou M, Harland L, Mustoe T, Williams S, Holdstock J, Yague E, et al. Transgenes encompassing dual-promoter CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. Genomics. 2003; 82:269–279. [PubMed: 12906852]

22. Williams S, Mustoe T, Mulcahy T, Griffiths M, Simpson D, Antoniou M, et al. CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnol. 2005; 5:17. [PubMed: 15935093]

23. Giraldo P, Montoliu L. Size matters: Use of YACs, BACs and PACs in transgenic animals. Transgenic Res. 2001; 10:83–103. [PubMed: 11305364]

24. Truffinet V, Guglielmi L, Cogné M, Denizot Y. The chicken β-globin HS4 insulator is not a silver bullet to obtain copy-number dependent expression of transgenes in stable B cell transfectants. Immunol Lett. 2005; 96:303–304. [PubMed: 15585338]

25. Blaas L, Musteanu M, Eferl R, Bauer A, Casanova E. Bacterial artificial chromosomes improve recombinant protein production in mammalian cells. BMC Biotechnol. 2009; 9:3. [PubMed: 19144171]

26. Bian Q, Belmont AS. BAC TG-EMBED: one-step method for high-level, copy-number-dependent, position-independent transgene expression. Nucleic Acids Res. 2010; 38:e127. [PubMed: 20385594]

27. Li Y, Liu W, Oo TF, Wang L, Tang Y, Jackson-Lewis V, et al. Mutant LRRK2(R1441G) BAC transgenic mice recapitulate cardinal features of Parkinson's disease. Nat Neurosci. 2009; 12:826–8. [PubMed: 19503083]

28. Zboray K, Sommeregger W, Bogner E, Gili A, Sterovsky T, Fauland K, et al. Heterologous protein production using euchromatin-containing expression vectors in mammalian cells. Nucleic Acids Res. 2015; 43:gkv475.

29. Hu Y, Plutz M, Belmont AS. Hsp70 gene association with nuclear speckles is Hsp70 promoter specific. J Cell Biol. 2010; 191:711–9. [PubMed: 21059845]

30. Bian Q, Khanna N, Alvikas J, Belmont AS. β-Globin cis-elements determine differential nuclear targeting through epigenetic modifications. J Cell Biol. 2013; 203:767–83. [PubMed: 24297746]

31. Sinclair P, Bian Q, Plutz M, Heard E, Belmont AS. Dynamic plasticity of large-scale chromatin structure revealed by self-assembly of engineered chromosome regions. J Cell Biol. 2010; 190:761–76. [PubMed: 20819934]

32. Johnson LF, Fuhrman CL, Wiedemann LM. Regulation of dihydrofolate reductase gene expression in mouse fibroblasts during the transition from the resting to growing state. J Cell Physiol. 1978; 97:397–406. [PubMed: 730776]

33. Meilinger D, Fellinger K, Bultmann S, Rothbauer U, Bonapace IM, Klinkert WEF, et al. Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. EMBO Rep. 2009; 10:1259–64. [PubMed: 19798101]

34. Cherry SR, Biniszkiewicz D, van Parijs L, Baltimore D, Jaenisch R. Retroviral expression in embryonic stem cells and hematopoietic stem cells. Mol Cell Biol. 2000; 20:7419–7426. [PubMed: 11003639]
35. Lois C, Hong EJ, Pease S, Brown EJ, Baltimore D. Germline Transmission and Tissue-Specific Expression of Transgenes Delivered by Lentiviral Vectors. Science. 2002; 295:868–872. [PubMed: 11786607]
36. Barber RD, Harmer DW, Coleman RA, Clark BJ. GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiol Genomics. 2005; 21:389–395. [PubMed: 15769908]
37. Zhao B, Chaturvedi P, Zimmerman DL, Andrew S. BAC TG-EMBED tools for either episomal or integrated transgene expression. Nucleic Acids Res. Manuscript submitted.
38. Chusainow J, Yang YS, Yeo JHM, Toh PC, Asvadi P, Wong NSC, et al. A study of monoclonal antibody-producing CHO cell lines: what makes a stable high producer? Biotechnol Bioeng. 2009; 102:1182–96. [PubMed: 18979540]
39. Dadheech N, Srivastava A, Belani M, Gupta S, Pal R, Bhonde RR, et al. Basal expression of pluripotency-associated genes can contribute to stemness property and differentiation potential. Stem Cells Dev. 2013; 22:1802–17. [PubMed: 23343006]
40. Shao HY, Hsu HY, Wu KS, Hee SW, Chuang LM, Yeh JI. Prolonged induction activates Cebpα independent adipogenesis in NIH/3T3 cells. PLoS One. 2013; 8:e51459. [PubMed: 23326314]
41. Bibel M, Richter J, Lacroix E, Barde YA. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. Nat Protoc. 2007; 2:1034–43. [PubMed: 17546008]
42. Katsetos CD, Legido A, Perentes E, Mörk SJ. Class III beta-tubulin isotype: a key cytoskeletal protein at the crossroads of developmental neurobiology and tumor neuropathology. J Child Neurol. 2003; 18:851–66. discussion 867. [PubMed: 14736079]
43. Yasunaga M, Tada S, Torikai-Nishikawa S, Nakano Y, Okada M, Jakt LM, et al. Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. Nat Biotechnol. 2005; 23:1542–50. [PubMed: 16311587]
44. Hong S, Hwang DD, Yoon S, Isacson O, Ramezani A, Hawley RG, et al. Functional Analysis of Various Promoters in Lentiviral Vectors at Different Stages of In Vitro Differentiation of Mouse Embryonic Stem Cells. Mol Ther. 2007; 15:1630–1639. [PubMed: 17609656]
45. Hamaguchi I, Woods NB, Panagopoulos I, Andersson E, Mikkola H, Fahlman C, et al. Lentivirus vector gene expression during ES cell-derived hematopoietic development *in vitro*. J Virol. 2000; 74:10778–10784. [PubMed: 11044122]
46. Kunert R, Casanova E. Recent advances in recombinant protein production: BAC-based expression vectors, the bigger the better. Bioengineered. 2013; 4:258–261. [PubMed: 23680894]
47. Lombardo A, Cesana D, Genovese P, Di Stefano B, Provasi E, Colombo DF, et al. Site-specific integration and tailoring of cassette design for sustainable gene transfer. Nat Methods. 2011; 8:861–869. [PubMed: 21857672]
48. Tasic B, Hippenmeyer S, Wang C, Gamboa M, Zong H, Chen-Tsai Y, et al. Site-specific integrase-mediated transgenesis in mice via pronuclear injection. Proc Natl Acad Sci USA. 2011; 108:7902–7907. [PubMed: 21464299]
49. Bronson SK, Plaehn EG, Kluckman KD, Hagaman JR, Maeda N, Smithies O. Single-copy transgenic mice with chosen-site integration. Proc Natl Acad Sci USA. 1996; 93:9067–9072. [PubMed: 8799155]
50. Giel-Moloney M, Krause DS, Chen G, Van Etten RA, Leiter AB. Ubiquitous and uniform in vivo fluorescence in ROSA26-EGFP BAC transgenic mice. Genesis. 2007; 45:83–9. [PubMed: 17269129]
51. Gong S, Zheng C, Doughty ML, Losos K, Didkovsky N, Schambra UB, et al. A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature. 2003; 425:917–925. [PubMed: 14586460]
52. Luo Y, Liu C, Cerbini T, San H. Stable enhanced green fluorescent protein expression after differentiation and transplantation of reporter human induced pluripotent stem cells generated by AAVS1 transcription. Stem Cells Transl Med. 2014; 3:821–835. [PubMed: 24833591]

53. Hyde SC, Pringle IA, Abdullah S, Lawton AE, Davies LA, Varathalingam A, et al. CpG-free plasmids confer reduced inflammation and sustained pulmonary gene expression. Nat Biotechnol. 2008; 26:549–51. [PubMed: 18438402]
54. Maison C, Almouzni G. HP1 and the dynamics of heterochromatin maintenance. Nat Rev Mol Cell Biol. 2004; 5:296–304. [PubMed: 15071554]
55. Chen ZY, Riu E, He CY, Xu H, Kay MA. Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. Mol Ther. 2008; 16:548–556. [PubMed: 18253155]
56. Dalle B, Rubin JE, Alkan O, Sukonnik T, Pasceri P, Yao S, et al. eGFP reporter genes silence LCRb-globin transgene expression via CpG dinucleotides. Mol Ther. 2005; 11:591–599. [PubMed: 15771961]
57. Ebisuya M, Yamamoto T, Nakajima M, Nishida E. Ripples from neighbouring transcription. Nat Cell Biol. 2008; 10:1106–13. [PubMed: 19160492]
58. Akhtar W, de Jong J, Pindyurin AV, Pagie L, Meuleman W, de Ridder J, et al. Chromatin position effects assayed by thousands of reporters integrated in parallel. Cell. 2013; 154:914–27. [PubMed: 23953119]
59. Warming S, Costantino N, Court DL, Jenkins NA, Copeland NG. Simple and highly efficient BAC recombineering using galK selection. Nucleic Acids Res. 2005; 33:e36. [PubMed: 15731329]
60. Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, et al. The UCSC Genome Browser Database. Nucleic Acids Res. 2003; 31:51–4. [PubMed: 12519945]
61. Khanna N, Bian Q, Plutz M, Belmont AS. BAC manipulations for making BAC transgene arrays. Methods Mol Biol. 2013; 1042:197–210. [PubMed: 23980009]
62. Strukov YG, Belmont AS. Development of mammalian cells lines with lac operator-tagged chromosomes. Cold Spring Harb Protoc. 2008; 3:1–9.
63. Faust D, Schmitt C, Oesch F, Oesch-Bartlomowicz B, Schreck I, Weiss C, et al. Differential p38-dependent signalling in response to cellular stress and mitogenic stimulation in fibroblasts. Cell Commun Signal. 2012; 10:6. [PubMed: 22404972]
64. Metzger JM, Lin WI, Samuelson LC. Vital staining of cardiac myocytes during embryonic stem cell cardiogenesis in vitro. Circ Res. 1996; 78:547–552. [PubMed: 8635211]
65. Agard DA, Hiraoka Y, Shaw P, Sedat JW. Fluorescence microscopy in three dimensions. Methods Cell Biol. 1989; 30:353–77. [PubMed: 2494418]

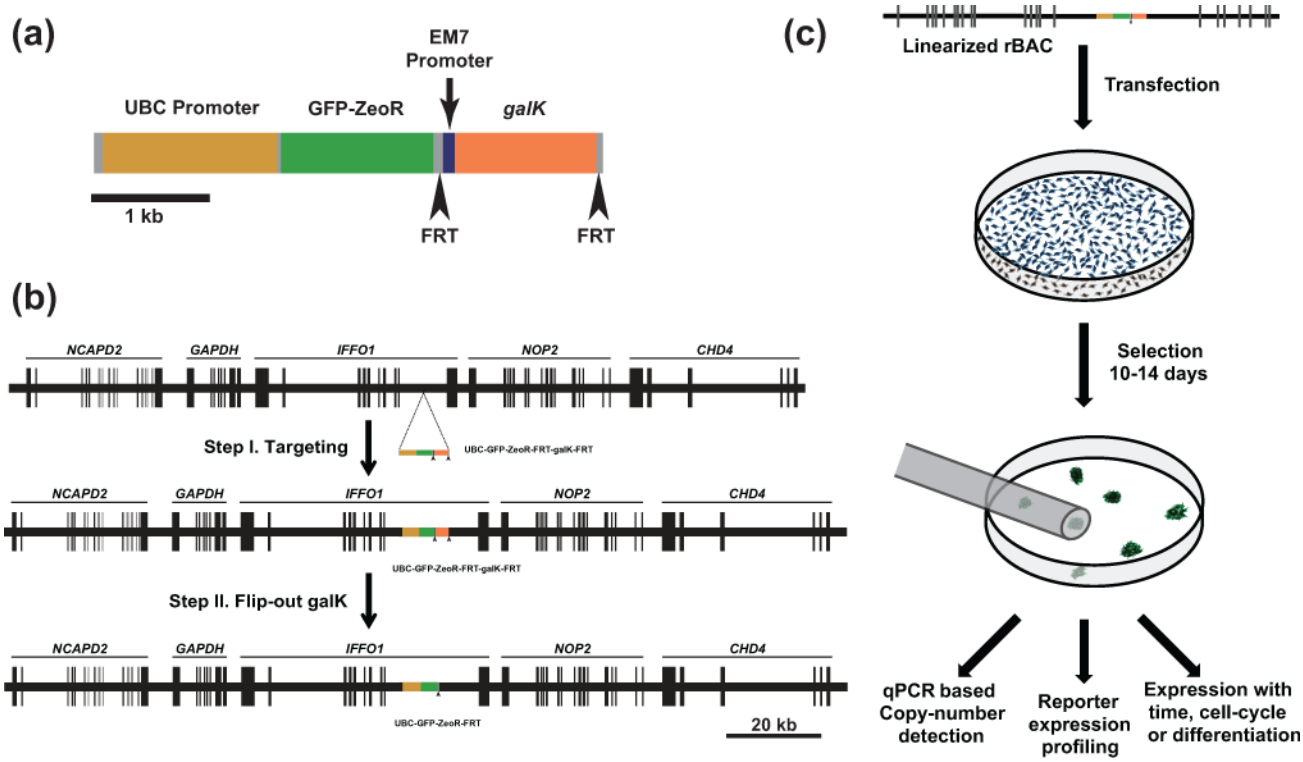

Figure 1.

Summary of experimental design. (a) UGG construct showing GFP-ZeoR minigene and the *GalK* positive/negative selection marker flanked by 34 bp flippase recognition target (FRT) sites (arrowheads). (b) Insertion of UBC-GFP-ZeoR construct into GAPDH BAC by BAC recombineering: Step I- homologous recombination of UGG construct into BAC intergenic region (dotted lines). Step 2- Flipping-out the *GalK* selection marker using *Flp* recombinase. (c) Overview of reporter expression assay: transfection of linearized BAC, selection of stable transformed colonies, and then qPCR determination of BAC copy number and measurement of reporter gene fluorescence. Reporter gene fluorescence of a subset of clones was monitored as a function of cell cycle arrest and/or cell differentiation.

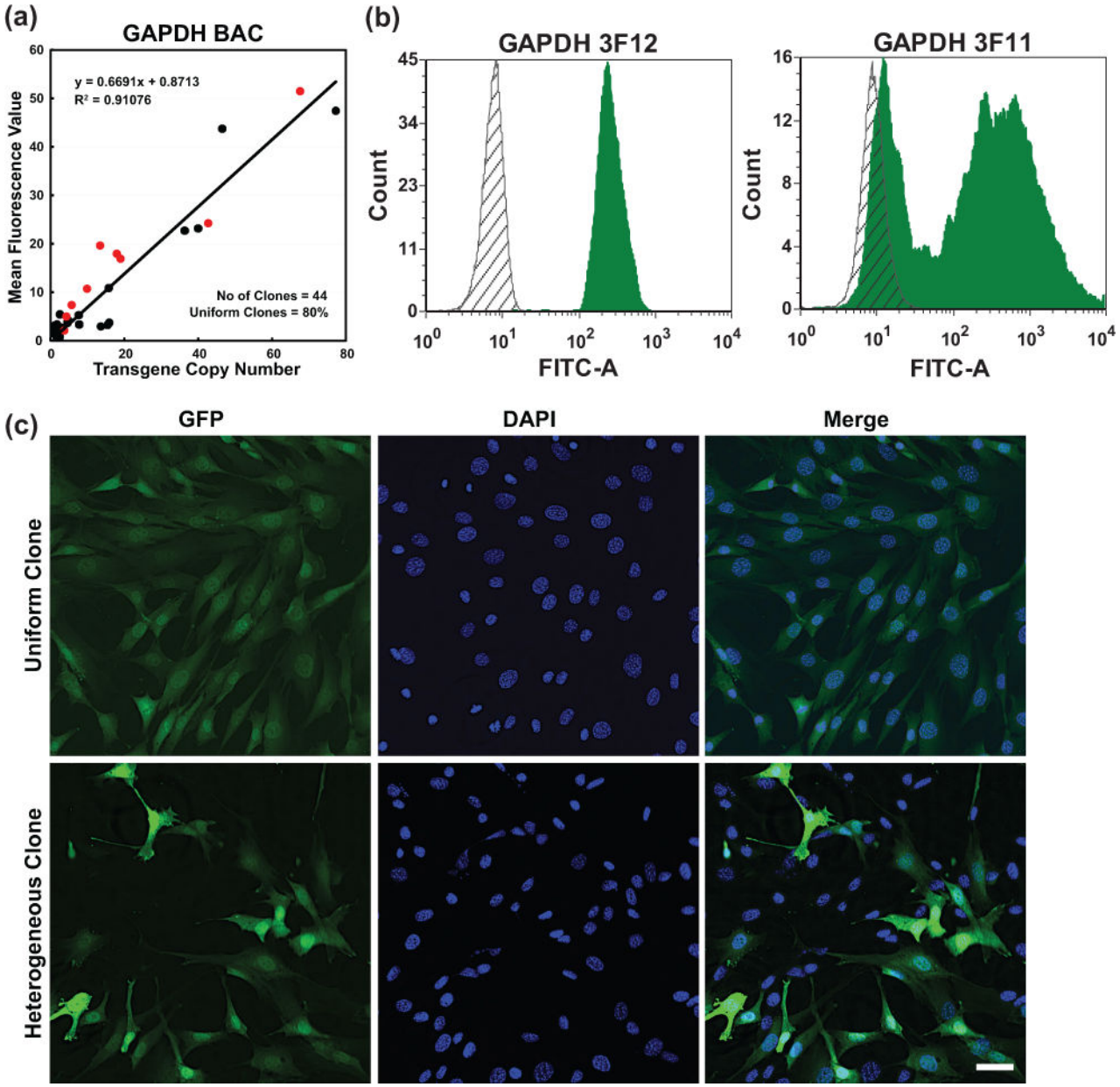

Figure 2.

Copy-number dependent expression but two types of reporter expression distributions. (a) Average GFP-ZeoR expression levels of individual clones normalized to reference fluorescent beads (y-axis) versus transgene copy-number (x-axis): linear regression fit (black line) with y-intercepts set to background fluorescence levels are shown with corresponding correlation coefficient and equation. Uniformly expressing clones (black circles), heterogeneously expressing clones (red circles). Bottom right- Number of clones analyzed and percentage of clones with uniform reporter expression. (b) Expression distributions of two stable NIH 3T3 clones carrying GAPDH BAC after 28 days of transfection with similar mean expression levels. Clone 3F12 shows uniform expression (left), while clone 3F11 shows heterogeneous expression (right). Autofluorescence of untransfected cells is plotted in

shaded gray. Fluorescence is measured in arbitrary units. (c) Fluorescence images of asynchronous cultures of the same uniform (top) and heterogeneous (bottom) NIH 3T3 clones. Nuclei are counterstained with DAPI (blue). Scale bar = 10 μm.

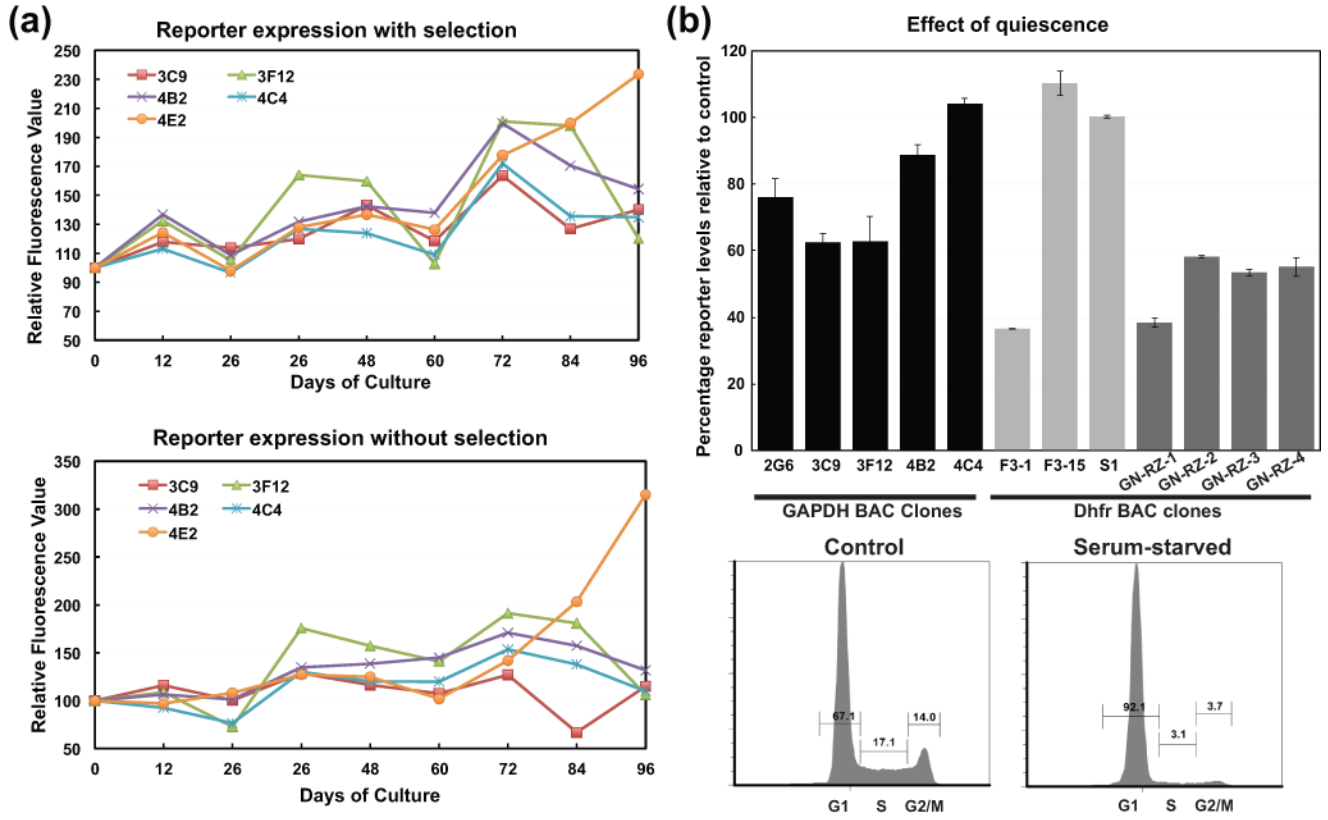

Figure 3.

Stability of BAC TG-EMBED UBC-GFP-ZeoR reporter expression in NIH 3T3 cell clones with and without selection and after cell-cycle arrest. (a) Reporter gene expression (y-axis, fluorescence levels normalized relative to day 0) for five uniformly expressing clones over 90 days of continuous passaging (x-axis, days of culture) with (top) or without (bottom) zeocin selection. (b) Minimal changes in reporter expression for 5 GAPDH BAC and 3 Dhfr BAC transgene clones carrying GFP-ZeoR reporter after low-serum G0 arrest. As a control 4 Dhfr BAC transgene clones carrying CMV-GFP reporter after low-serum G0 arrest are shown.[26] Top: Relative reporter expression levels (y-axis) normalized relative to asynchronous cultures (n=3, mean ± SEM). Bottom: Representative flow cytometry profiles of propidium iodide (PI) stained control (left) and G0-arrested (right) cells. Percentages of cells with 2N (G1), intermediate (S), or 4N (G2/M) DNA content estimated by fitting these profiles are shown in black text.

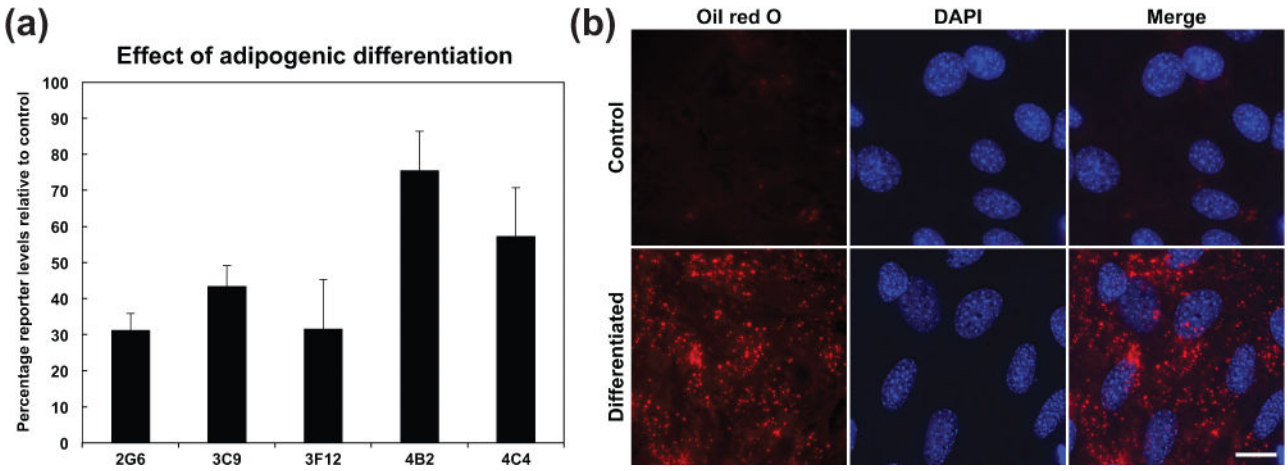

Figure 4.

Reporter expression is maintained after adipocyte differentiation of NIH 3T3 clones. (a) Percentage of pre-differentiation reporter gene mean GFP fluorescence measured in 5 NIH 3T3 cell clones with independent BAC transgenes integrations after differentiation into adipocytes (n=3, mean ± SEM). Fluorescence levels represent the mean from flow cytometry measurements. (b) Representative micrographs of Oil Red O (red) stained control and transdifferentiated cells. DNA stained with DAPI (blue). Scale bar = 5 μm.

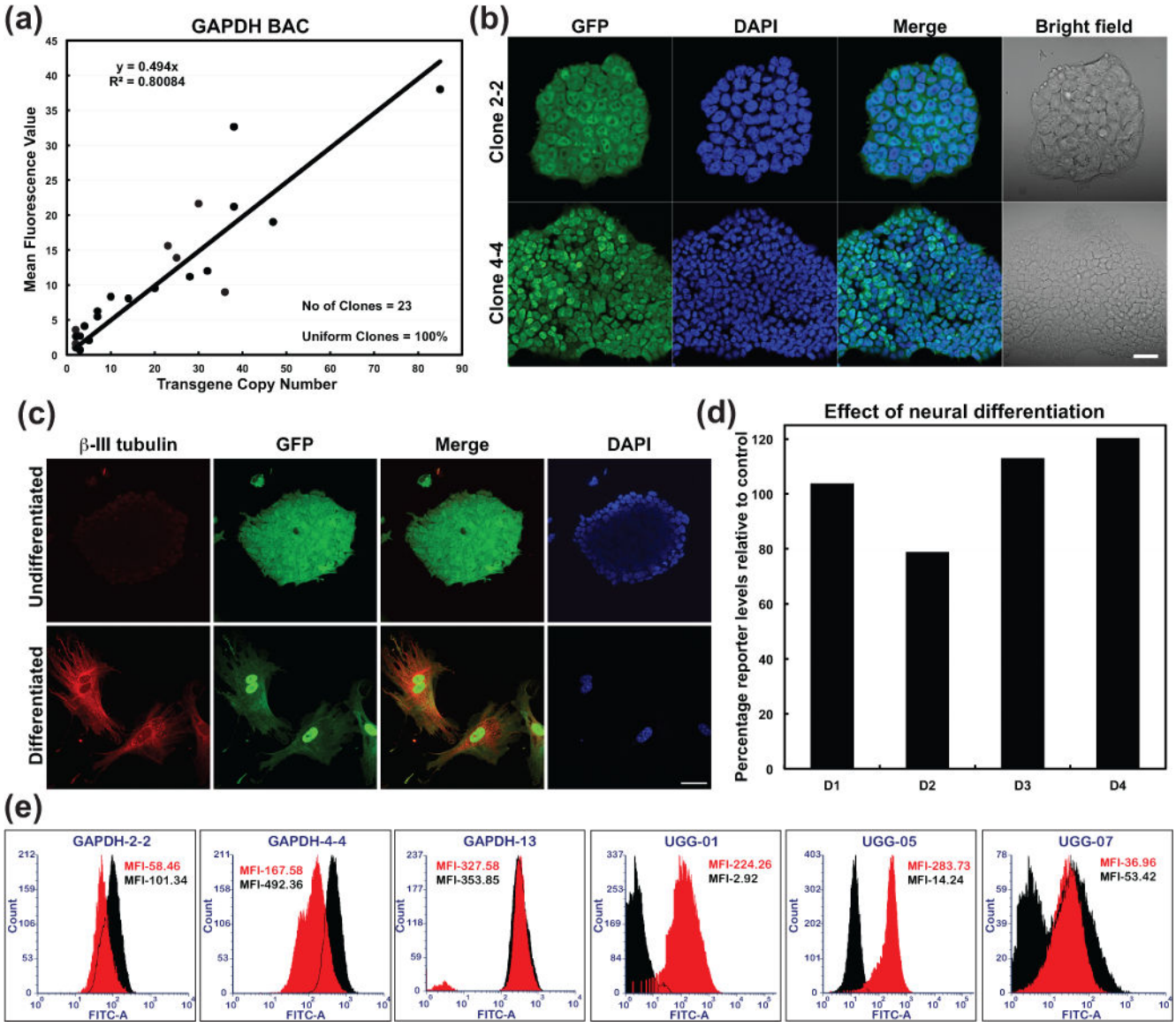

Figure 5.

Uniform BAC TG-EMBED reporter expression in mouse undifferentiated ESC clones is maintained after differentiation into neuronal precursors. (a) Copy-number dependent expression of BAC reporter genes. Average normalized GFP-ZeoR expression levels (y-axis) versus transgene copy-number (x-axis) for 23 stably transfected cell clones. Linear regression fit (black line) with y-intercepts set at background fluorescence levels are shown with corresponding correlation coefficients and equations. (b) Uniform GFP-ZeoR reporter fluorescence (green) throughout undifferentiated ESC colonies from two undifferentiated HM1 mouse ES cell transgenic clones. Nuclei are counterstained with DAPI (blue). (c) GFP-ZeoR reporter expression (green) is maintained after ESC differentiation into neuronal precursors- undifferentiated ESCs (top row) versus ESCs differentiated into neuronal precursors (bottom row), with panneural marker β-III tubulin immunostaining (red). Nuclei are counterstained with DAPI (blue). (d) Percentage change in GFP-ZeoR reporter

expression relative to the undifferentiated control for four independent neural differentiation experiments with ES cell clone 13 carrying GAPDH BAC. (e) Expression distributions of stable ES cell clones carrying GAPDH BAC (Clones 2-2, 4-4 and 13) or UGG reporter plasmid (Clones 01, 05 and 07) after neural differentiation. GFP-ZeoR reporter expression of undifferentiated ES cells (red) and corresponding differentiated cells (black) are plotted. Mean fluorescence (MFI) of undifferentiated ES cells (red) and differentiated cells (black) is measured in arbitrary units and indicated on respective plots. (b&c) Scale bar = 20 μm.

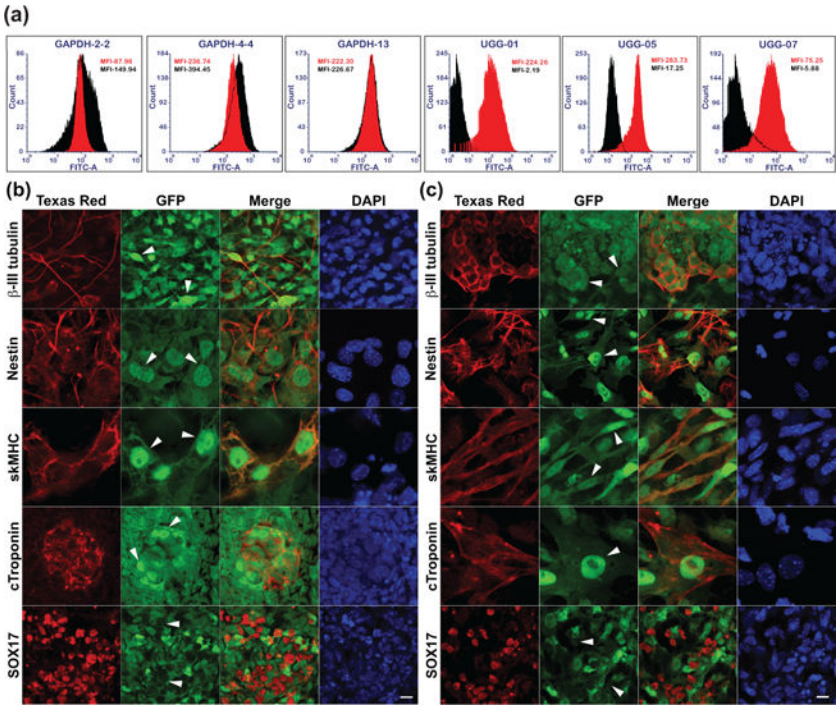

Figure 6.

BAC TG-EMBED reporter expression is maintained within most differentiated cells from all three germ layers in embryoid bodies derived from undifferentiated mESCs transfected with BAC transgenes. (a) Expression distributions of stable ES cell clones carrying GAPDH BAC (Clones 2-2, 4-4 and 13) or UGG reporter plasmid (Clones 01, 05 and 07). GFP-ZeoR reporter expression of undifferentiated ES cells (red) and corresponding spontaneously differentiated cells (black) are plotted. Mean fluorescence (MFI) of undifferentiated ES cells (red) and differentiated cells (black) is measured in arbitrary units and indicated on respective plots. (b-c) Immunostaining panels showing GFP-ZeoR reporter expression (green) in different types of differentiated cells identified by different lineage specific markers (red) as indicated. Undifferentiated mESCs from two different HM1 mouse ES cell clones (b-Clone 2-2, c-Clone 4-4) containing integrated copies of the GAPDH BAC with UBC-GFP-ZeoR reporter gene were used to form *in vitro*, spontaneously differentiated embryoid bodies. Arrowheads show cells positively stained for respective lineage markers. Nuclei are counterstained with DAPI (blue). Scale bar = 20 μm. The data is representative of 3 independent experiments.