

Homogeneity and persistence of transgene expression by omitting antibiotic selection in cell line isolation

Weimin L. Kaufman, Ibrahim Kocman, Vishal Agrawal, Hans-Peter Rahn, Daniel Besser and Manfred Gossen*

Max Delbrück Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany

Received April 30, 2008; Revised July 10, 2008; Accepted July 24, 2008

ABSTRACT

Nonuniform, mosaic expression patterns of transgenes are often linked to transcriptional silencing, triggered by epigenetic modifications of the exogenous DNA. Such phenotypes are common phenomena in genetically engineered cells and organisms. They are widely attributed to features of transgenic transcription units distinct from endogenous genes, rendering them particularly susceptible to epigenetic downregulation. Contrary to this assumption we show that the method used for the isolation of stably transfected cells has the most profound impact on transgene expression patterns. Standard antibiotic selection was directly compared to cell sorting for the establishment of stable cells. Only the latter procedure could warrant a high degree of uniformity and stability in gene expression. Marker genes useful for the essential cell sorting step encode mostly fluorescent proteins. However, by combining this approach with site-specific recombination, it can be applied to isolate stable cell lines with the desired expression characteristics for any gene of interest.

INTRODUCTION

The genomic integration of recombinant transcription units is the standard technology for the genetic manipulation of mammalian cells, both in basic research applications as well as in biotechnology (1,2). Once a stably transfected cell clone has been identified, it should ideally express the transgene in all cells, and at constant levels over prolonged periods of time. Vastly different expression levels are frequently observed between individual clones.

These differences are attributed to either copy number effects or, more importantly, the particular chromosomal integration site. Within a given clone, cell-to-cell variations in expression levels (also referred to as mosaicism or variegation) as well as complete cessation of gene expression over time is quite common, despite the physical integrity of the newly acquired transgene. In most cases the loss of overall expression levels, or silencing, can be easily monitored. In contrast, heterogeneity in expression often goes unnoticed, either due to unawareness of the problem or the lack of appropriate *in situ* tools for the detection of mosaicism. The same problem applies to transgenic animals, where mosaic expression patterns are also common. The interpretation of such variegated expression patterns is greatly aided by the availability of dynamic live cell reporter systems, allowing to repeatedly monitor gene expression of a cell population, both quantitatively and qualitatively. In comparison, a single 'snapshot' analysis revealing a heterogeneous expression pattern can result either from (i) the asynchronous conversion of an expressing to a nonexpressing cell population by gene silencing, (ii) intrinsic fluctuations of gene expression or (iii) a rather stable distribution of expressing and nonexpressing cells. Irrespective of the underlying process, variegation is often an obstacle for the application of stable cell lines (3,4).

Progressive silencing and the generation of mosaic expression patterns are often caused by epigenetic downregulation of transgenes. This process manifests itself via DNA methylation and by locally restricted, covalent modifications of histones, resulting in the formation of repressive chromatin structures (5,6). The susceptibility of exogenous transcription units to dynamic changes in their epigenetic state is frequently ascribed to characteristics of transgenes that distinguish them from endogenous transcription units. Among the factors considered are (i) the frequently observed high copy number of transgenes and the resulting tandem array integration pattern,

*To whom correspondence should be addressed. Tel: +49 30 94062596; Fax: +49 30 94062551; Email: mgossen@mdc-berlin.de
Present address:

Weimin L. Kaufman, Center for Pediatric Biomedical Research, University of Rochester, 601 Elmwood Ave., Rochester, NY 14642, USA

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

contributing to repeat-induced silencing, (ii) the nature of the expression signals employed for transgenesis, especially the promoter choice and (iii) particular features of the transgene itself, which can reside in the nucleotide composition and sequence of the DNA transferred, like the frequency and patterns of CpG dinucleotides. These particularities of transgenes have been implied both in variegated as well as in overall gene silencing (7–10). In addition, it is also possible that a response of the host cell is triggered by the actual DNA integration process itself. This could result in the localized disruption of an established chromatin context irrespective of particular characteristics of the incoming DNA. It should be noted, however, that mosaic expression patterns are not restricted to transgenes, but have also been documented for endogenous genes. For a given gene analyzed, the cause for such phenotypic variegation could be heterogeneity in its local epigenetic environment within a cell population, stochastic events contributing to its expression or dynamic expression control mechanisms resulting in the oscillation of its activity status (11–15). A more detailed understanding of the interplay of these factors is not only of general interest for the regulation of genome-wide transcription programs, it may also provide us with a blueprint for the design of transgenic transcription units and transgenesis protocols less prone to undesired epigenetic modulation.

To this end, we initially set out to identify features of transgenic transcription units that facilitate their homogeneous expression in stable cell lines. However, we found that the method of enriching and isolating transgene-positive cells had the greatest impact on homogeneity and stability of gene expression. When compared directly to a standard antibiotic selection regimen, cell populations isolated by fluorescence-activated cell sorting (FACS) on the basis of fluorescent protein expression showed little cell-to-cell variation and the high levels of transgene expression were remarkably stable over time. By using site-specific recombinases, we were able to transfer these qualitative expression parameters to other transgenes, bypassing antibiotic selection protocols with their bias towards mosaic expression patterns.

MATERIALS AND METHODS

Plasmid constructs

Plasmid pCMV-EGFP is a modified version of pEGFP-C1 (Clontech, Mountain View, CA, USA), which was used in initial experiments. Both vectors carrying a neomycin antibiotic resistance marker gene under the control of an SV40 promoter. Replacement of the CMV promoter in pCMV-EGFP by the promoter of the human elongation factor 1 α [EF1 α from pEF1; (16)] resulted in pEF-EGFP. In both plasmids the reporter cassette (promoter/enhancer—EGFP—polyA signal) can be precisely excised. For some of the experiments the plasmids were linearized in the vector backbone. Plasmids allowing for the establishment of stable cell lines after sorting for fluorescent protein expression and recombinase-mediated deletion of the marker gene do not contain an antibiotic resistance marker gene for selection in mammalian cells. They are

based on a plasmid where EGFP, with its own polyadenylation signal, is flanked by two FLP recombinase targets of the F3 type (17) in the same orientation. Expression is controlled by EF1 α . The promoter-distal F3 site is followed by a multiple cloning site (mcs) and the bovine growth hormone polyadenylation signal (pEF_{F3}EGFP_{F3}mcs; unique sites of the mcs, promoter proximal to promoter distal: EcoRI, XmaI, BamHI, SpeI, NotI, XhoI, NsiI). The mcs was used to insert cDNAs of the monomeric red fluorescent protein [mRFP (18)] and the human transcription factor Oct4, resulting in pEF_{F3}EGFP_{F3}mRFP and pEF_{F3}EGFP_{F3}Oct4, respectively. Details of the plasmid construction (including the complete vector sequences) are available upon request.

pCAGGS-FLPe was used for the transient expression of an enhanced version of the FLP recombinase (19).

Cell culture and transfections

Cells were cultured at 37°C with 5% CO₂ in minimal essential medium containing 10% bovine serum, 2 mM L-glutamine, 50 μ g/ml penicillin and 50 μ g/ml streptomycin. Doubling time was about 21 h for HeLa cells and about 19 h for CHO cells. G418 (Invitrogen, Carlsbad, CA, USA) was used at a concentration of 400 μ g/ml. Subcloning by limited dilution was achieved by seeding cells in 96-well plates at a density of 0.2 cells/well. For both transient and stable transfections the Roti[®]-Fect liposome formulation (Carl Roth GmbH, Karlsruhe, Germany) was used according to the manufacturer's protocol.

Flow cytometry and preparative FACS

Exponentially growing cells were harvested and resuspended in PBS supplemented with 0.1% EDTA. Fluorescent protein expression was analyzed on a FACSCalibur (BD Biosciences, San Jose, CA, USA) using argon-ion laser excitation (488 nm). EGFP was detected using the FL1 parameter (emission filter: 530 \pm 15 nm) and mRFP by using the FL2 parameter (emission filter 585 \pm 21 nm). Dead cells and debris were excluded from the analysis by gating for intact cells using the forward and sideward scatter. Untransfected cells were always included as control to detect autofluorescence. Data acquisition was carried out by analyzing 10 000 events/sample using CellQuest Software (BD Biosciences). To enrich EGFP-expressing cells, cells were sorted on a FACS Vantage SE (BD Biosciences, Carlsbad, CA, USA) with fluorescent parameters identical to those described above. Cells were sorted directly into growth media. FACS data were analyzed using FlowJo Software (Tree Star, Inc., Ashland, OR, USA).

Immunoblot and immunofluorescence analysis

Immunoblot analysis for Oct4 expression was by standard methods, using a goat antiserum directed against human Oct3/4 (dilution 1 : 1000; R&D Systems, Minneapolis, MN, USA) and a mouse monoclonal antibody directed against α -tubulin (dilution 1 : 4000; Ab-1, Oncogene Science, Cambridge, MA, USA) as a loading control. Signal detection was by enhanced chemiluminescence. The same antibodies (anti-Oct3/4 at 1 : 200; anti- α -tubulin at 1 : 400) were used for immunofluorescence analysis of cells grown in chamber slides (Lab-Tek #136439; Nalge Nunc, Naperville,

IL, USA) and treated with 4% paraformaldehyde/0.1% Triton. Cy3-conjugated donkey anti-goat antibody and Cy2-conjugated donkey anti-mouse antibody (both 1:400; Jackson ImmunoResearch, West Grove, PA, USA) were used as secondary reagents. Images were taken on a DMI6000 B microscope (Leica, Wetzlar, Germany) using a 10× objective and a DFC 350 FX camera (Leica). Fluorescence was detected with filters L5 and N2.1. Images were processed by using the LAS AF application suite 1.6.2 (Leica). Image overlays were created using Adobe Photoshop, version 7.0. For quantification of EGFP expression in CHO cells, we analyzed total cell extracts from known numbers of cells by immunoblotting, in parallel with purified EGFP protein (MBL, Woburn, MA, USA). Detection was by using a rabbit antiserum against GFP (Invitrogen).

DNA and RNA analysis

Southern blotting was performed as described (20). For real-time RT-PCR analysis, total RNA was isolated using Trizol (Invitrogen) from HeLa cells and the respective cell lines stably transfected with pEF_{F3}EGFP_{F3}hOct4 before and after FLP-recombination. The RNA was reverse transcribed using Moloney murine leukemia virus reverse transcriptase (Promega, Madison, WI, USA) according to the manufacturer's instructions. Real-time RT-PCR was performed on the iCycler IQ TM 5 multicolor real-time detection system (Bio-Rad, Hercules, CA, USA) according to standard protocols, using ABsolute™ SYBR green fluorescein (ABgene, Surrey, UK).

RESULTS

Isolation of stably transfected cells by FACS

To facilitate a direct comparison between antibiotic selection and FACS for the isolation of cells expressing a transgene after chromosomal integration, we first analyzed two cell populations derived from a single transfection experiment. We used a plasmid harboring both a CMV promoter driving expression of the green fluorescent protein (GFP), as well as a transcription unit conferring neomycin resistance. After transfection of the linearized vector into HeLa cells, its dual function enabled antibiotic selection of G418-resistant cells and cell sorting for fluorescent protein expression in parallel experiments. FACS enrichment of GFP-positive cells was done twice and the fluorescence level of the resulting cell population compared directly to the GFP expression of pooled G418-resistant clones originating from the same transfection. Figure 1 shows that contrary to the sorted cell pool, the vast majority of G418-resistant cells showed little to no GFP signal. Microscopic examination of G418-resistant clones prior to pooling had shown that the few GFP-positive cells did not originate from a rare clone with intense GFP expression, but that a low percentage of positive cells was visible in several individual colonies. The high level of GFP expression in the sorted cell pool showed only a minor decrease when reanalyzed after growing those cells for about 30 population doublings (PDs) without applying selective pressure through addition of G418 (Figure 1). All subsequent cell sorting experiments were done according

to the basic flow chart outlined in Figure 1. We did not notice any marked difference in the results when making minor adjustments to the parameters indicated.

Characterization of clonal cell lines

Next we analyzed clonal cell lines, derived from cell sorted, GFP-positive pools by limited dilution. Their mean fluorescence values were compared with randomly chosen clones obtained by standard antibiotic selection. The results were in line with the analysis of the cell pools, with G418-selected clones being either completely GFP negative, or showing a weak GFP signal, often originating from a limited number of positive cells. In contrast, subcloned isolates from the FACS protocol were all positive for GFP expression according to quantitative flow cytometry (Figure 2A), and the majority of these clones was also unambiguously positive by fluorescence microscopy. We thus analyzed clones derived by the FACS protocol from cells transfected with a linearized pCMV-EGFP vector for their homogeneity according to microscopic inspection. More than one-third of the clones isolated were judged to be uniformly GFP positive (Figure 2B, left panel). In contrast, using the antibiotic selection protocols, we were never able to isolate uniformly expressing clones from stably transfected HeLa cells [see also (21) for comparison]. The ability to recover nonvariegated cell clones with high confidence allowed us to address various parameters frequently discussed to have a major impact on the formation of mosaic expression patterns. First, we asked if bacterial vector sequences have a major influence on transgene expression characteristics. To this end we compared the results obtained for the linearized pCMV-EGFP vector with that from a CMV-EGFP expression cassette. The isolated CMV-EGFP transcription unit (i.e. promoter, reporter gene and polyadenylation signal) devoid of bacterial vector sequences was transfected and cell lines were derived after FACS and subcloning by limited dilution. Figure 2B, middle panel, shows that a higher percentage of homogenous GFP expressing clones could be recovered. Second, the efficiency of obtaining such clones was further improved by using a corresponding EF-EGFP cassette, where the viral CMV promoter was exchanged for the cellular EF1 α promoter. This yielded about 80% uniformly transgene-positive clones (Figure 2B, right panel). Some of these latter clones were tested for the uniform expression persistence. They proved to be remarkably stable over >100 PDs (Figure 2C, upper panel), similar to clones obtained after transfection of the CMV-EGFP cassette (Figure 2C, middle panel). Note that these cells do not contain a selectable resistance marker. Thus, both CMV and EF1 α promoters can result in production of homogenous, long-term expressing clones, although the EF1 α promoter has a higher efficiency. Interestingly, clones expressing the GFP gene under EF1 α control showed a higher uniformity in the fluorescence signal as compared to CMV-driven GFP-positive cell clones. The same results for stability and homogeneity of transgene expression were found when the experiments were repeated with CHO cells using the linearized pCMV-EGFP

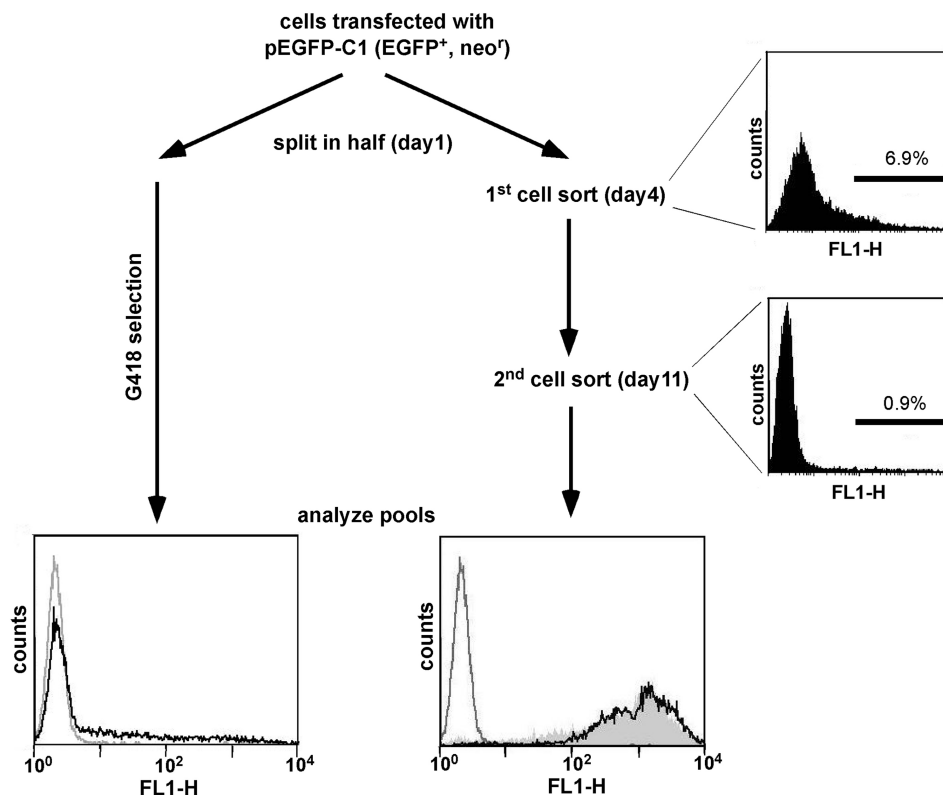


Figure 1. Flow chart of FACS and antibiotic-selection protocols. HeLa cells were transfected with pEGFP-C1. One day posttransfection (d.p.t.) the cells were divided, with one aliquot subjected to G418 selection to obtain a stably transfected cell population. The other aliquot of cells was grown in the absence of antibiotic selection, and cells were initially subjected to cell sorting for EGFP expressers at 4 d.p.t. The sorting gate and percentage of sorted cells of the whole population is indicated in the right inset showing the FACS profile for the fluorescence signal. Due to the transient nature of EGFP expression at this time point, the majority of the gated population lost the reporter gene expression when grown without antibiotic selection until the second cell sorting at 11 d.p.t. The residual positive population was expanded and analyzed for fluorescent protein expression in parallel with the population grown under drug selection, obtained by pooling all G418-resistant colonies. The FACS profiles at the bottom show the fluorescence signals from the pools of transfected cells (black line) in comparison to untransfected, parental HeLa cells (grey line). Further growth of the sorted population for 30 population doublings without selective pressure resulted only in a minor loss of the EGFP signal (grey-filled area in bottom right FACS profile).

vector (Figure 2C, lower panel). To get a better idea about the absolute expression levels observed, we determined in one of the three CHO clones analyzed the EGFP levels by quantitative immunoblot analysis. In this clone (indicated in Figure 2C, lower panel) the level of EGFP was about 0.5 pg per cell (about 10^7 molecules/cell).

In addition to the use of FACS for the isolation of stable cell lines, the same results can be achieved by using magnetic cell separation (MACS). We used a plasmid with CMV promoter driving expression of GFP, an internal ribosome entry site, and a cell surface marker gene, transcribed as a bicistronic mRNA [derived from MACSelect™ LNGFR System, Miltenyi Biotec, Bergisch Gladbach, Germany; see also (22)]. Magnetic isolation of transfected CHO cells was specific for expression of the cell surface receptor, and was performed in two rounds prior to cloning by limited dilution. Again, this procedure resulted in a high proportion of clones uniformly and stably expressing GFP, which was only used as an analytical marker, not for cell sorting (W.L.K. and M.G., data not shown). This demonstrates that the principle of stable cell line generation we describe is not limited by the use of FACS. However, in our hands the FACS approach was more robust than MACS and thus was pursued further.

Cell sorting as a general tool in establishing stable cell lines

We next asked if it is possible to adapt the approach outlined here for the generation of stable cell lines for transgenes that do not convey a sortable cellular phenotype. To this end, we constructed expression units as outlined in Figure 3A. Here, EF1 α drives expression of a GFP gene flanked by FRT sites arranged in the same orientation. Upon FLP recombination, the promoter proximal GFP gene and its polyadenylation signal will be deleted and a distal, previously not transcribed gene of interest would be placed under control of the EF promoter. To test this principle in an experimental setting that would allow us to directly monitor the succession of events, we initially chose a red fluorescent protein (RFP) gene. Cells were transfected and sorted for GFP expression and individual clones characterized for homogeneity and persistence of the GFP signal. These RFP-negative cells were transfected with a vector encoding enhanced FLP (19). Recombinase expression caused deletion of GFP and placement of RFP under EF1 α control, as indicated by previously green cells turning to red (Figure 3B). The deletion event was also monitored by Southern blotting (Figure 3C). After recloning, all of the

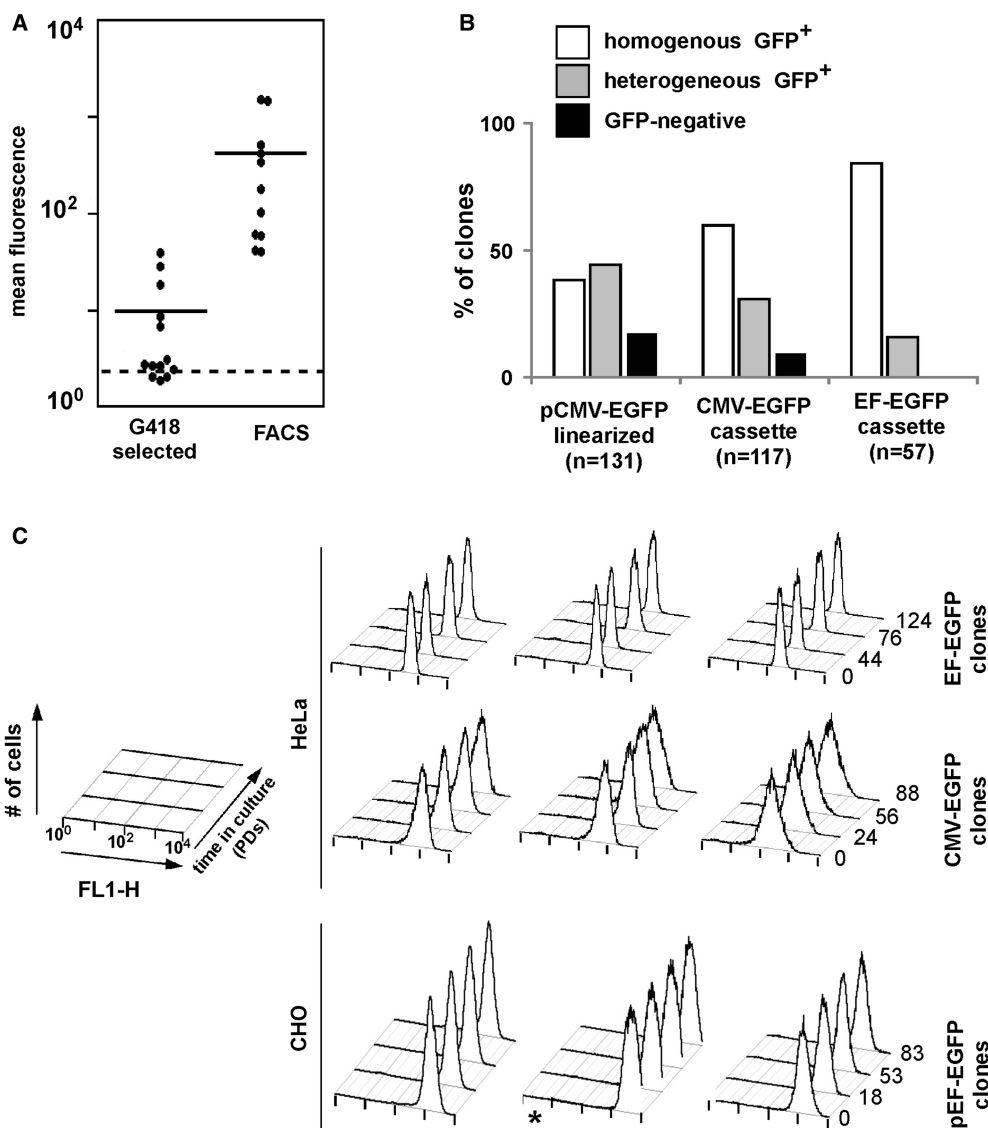


Figure 2. Characterization of clonal cell lines. Cell pools isolated by FACS were cloned by limited dilution and GFP expression patterns analyzed by flow cytometry and fluorescence microscopy. (A) Comparison of GFP expression levels (mean fluorescence) of randomly chosen clones, derived either from G418 selection or FACS as outlined in Figure 1. The background autofluorescence of untransfected cells is indicated by a dashed line. (B) Analysis of transgene expression patterns (homogeneously positive/heterogeneously positive/negative) according to microscopic analysis of GFP signals. The left panel shows the results for a linearized pCMV-EGFP vector, the middle panel for a corresponding expression cassette devoid of all plasmid backbone sequences. The right panel shows the results for such an expression cassette incorporating the EF promoter instead of the CMV promoter. The sample numbers of individual clones analyzed is indicated. (C) Clones derived from transfections with either EF or CMV expression units as indicated and that were judged homogeneously positive by microscopy were analyzed for homogeneity of transgene expression over time. The different time points of flow cytometric GFP profiling during continuous culturing of three independent clones for each experiment are indicated. The upper two panels depict HeLa cell clones, the lower panel CHO cell clones, with the one labeled by an asterisk used for quantitative immunoblotting.

GFP-negative clones analyzed were homogeneously RFP-positive, and maintained this expression pattern (Figure 3D).

Having shown the proof of principle, we turned to the expression of human oct4, the homogenous expression of which is of immediate relevance for the ongoing work in our laboratory. The human oct 4 gene was placed behind the FRT-flanked GFP cassette. Generation of stable, homogenous GFP-positive HeLa cell clones with a non-transcribed oct4 gene was as described before. HeLa cells were chosen as they are otherwise negative for oct4 expression. FLP-mediated deletion of the reporter cassette

caused a decrease in the GFP signal of the transfected subpopulation (Figure 4A). Subsequent recloning resulted in transgene expression in 2 out of 3 subclones analyzed, as shown by immunoblot analysis for Oct4 (Figure 4B). We corroborated these results by quantitative analysis of the oct4 transcript levels. cDNA obtained from total RNA was analyzed by qPCR. As shown in Figure 4C, the oct4 RNA levels of HeLa cells stably transfected with pEF_{F3}EGFP_{F3}Oct4 was barely above the low levels found in the parental HeLa cells. FLP-mediated cassette deletion resulted in a strong signal for the oct4 transcript. Lastly, the homogeneity of transgene expression was

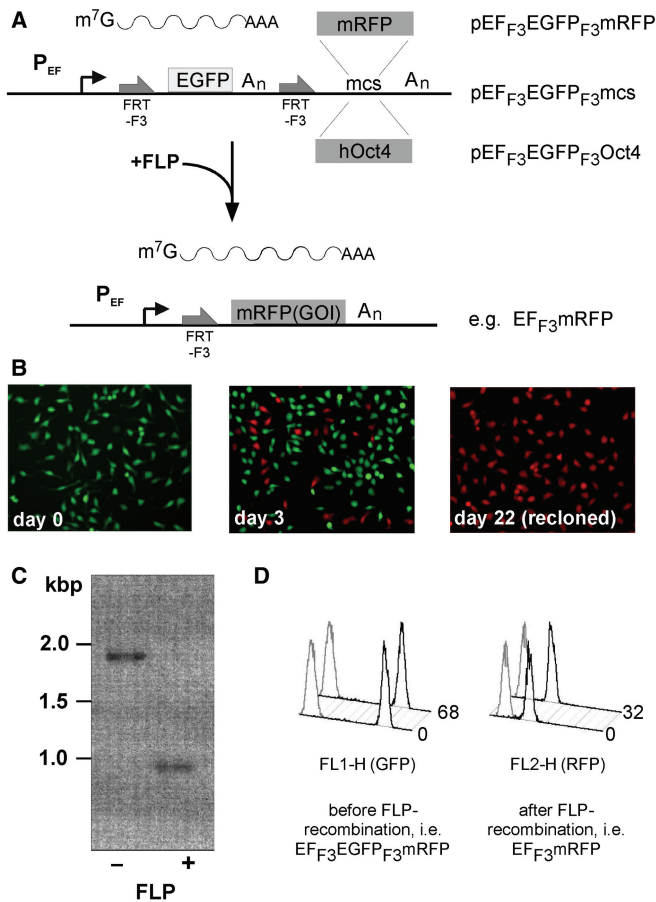


Figure 3. Recombinase-mediated switch of transgene expression. (A) Schematic presentation of the expression vectors for FLP-mediated switch of transgene expression used in this study. Genes of interest (GOI) inserted by us in the multiple cloning sites (mcs) were oct4 and mRFP. The deletion principle is outlined for the latter transgene, with the transcripts before and after recombination indicated. (B) Cells stably transfected with pEF_{F3}EGFP_{F3}mRFP were followed over the time course of a deletion experiment, showing both the GFP and RFP signals. Left panel: only EGFP was expressed. Middle panel: in cells where FLP levels upon transient transfection were sufficient to promote EGFP deletion, mRFP expression was activated. Untransfected cells stayed green. Right panel: recloning after FLP expression showed that GFP-negative clones were uniformly positive for RFP. (C) Southern blot analysis of a selected pEF_{F3}EGFP_{F3}mRFP clone before FLP-mediated recombination and a corresponding subclone after deletion of the EGFP reporter. The hybridization probe used is the mRFP gene. Note that the signal intensity for the mRFP target sequence decreased twofold after recombination (phosphorimager signal 246 versus 137 arbitrary units). The recombination principle anticipates a reduction of inserts to single copy transgenes. (D) The switch from GFP to RFP expression as monitored by flow cytometry.

monitored by immunofluorescence analysis for Oct4 protein. Prior to deletion of the GFP cassette, HeLa cells stably transfected with pEF_{F3}EGFP_{F3}Oct4 did not show a discernable Oct4 signal, in line with the immunoblot analysis. In contrast, after FLP-mediated deletion of the reporter cassette, clonal cells were uniformly positive for Oct4, with the transcription factor accumulating in the nucleus (Figure 4D).

We conclude from this series of experiments that qualitative parameters as identified for the expression of

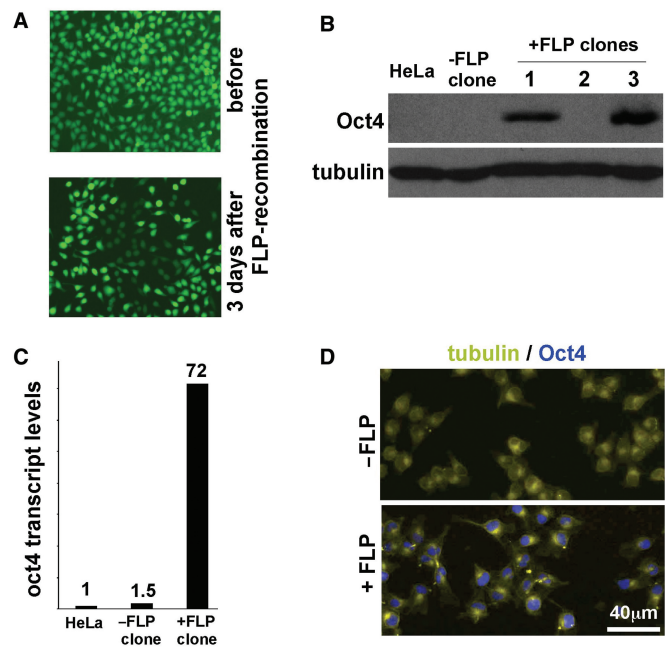


Figure 4. Homogenous expression of Oct4. (A) A homogeneously GFP-positive pEF_{F3}EGFP_{F3}hOct4 clone is shown in the top panel. Three days after transfection of a FLP expression vector a bimodal GFP pattern can be recognized (bottom panel), as FLP expressing cells show a progressive loss of the GFP signal upon reporter gene deletion. (B) Immunoblot analysis for Oct4 expression. Neither the non-recombined pEF_{F3}EGFP_{F3}hOct4 clone nor the recombined pEF_{F3}EGFP_{F3}hOct4 clone expressed detectable levels of Oct4. Of the three subclones isolated after FLP-mediated EGFP deletion, two showed a strong Oct4 signal. (C) RNA isolated from the parental cell line, the nonrecombined and a recombined pEF_{F3}EGFP_{F3}hOct4 clone was analyzed for Oct4-specific transcript levels by qPCR. (D) After FLP-mediated recombination, Oct4 (shown in blue) can be found in every nucleus of a pEF_{F3}EGFP_{F3}hOct4 subclone (lower panel), contrary to the initial, nonrecombined pEF_{F3}EGFP_{F3}hOct4 clone (upper panel) that was Oct4 negative. The counter stain in this immunofluorescence analysis was tubulin, shown in yellow.

a reporter gene can be transferred to a physically linked gene of interest by the experimental approach outlined.

DISCUSSION

The difference between ‘stable cell lines’ and ‘stable stable cell lines’ has been pointed out previously (23). This view of recombinant clonal cell populations and their tendency to silence transgene expression over time refers to the loss of net expression levels from engineered transgenes. While such a silencing might be gradual but homogenous, it can as well be restricted to a subpopulation among the cells. In general, such phenomena are attributed to epigenetic modulations of transgene activity in tissue cultures and transgenic animals. The interpretation of mosaic expression patterns in animals is complicated by the possibility that variegation might be caused by differentiation events, resulting in tissue microheterogeneity not uncovered by obvious histological changes. We, therefore, focused on analyzing transgene silencing in stably transfected cell lines, where a better understanding of this process might help us to bypass it. Our previous results showed that the

single-copy integration of transgenes into predetermined chromosome loci did not ensure identical phenotypic characteristics of the isogenic cell lines generated. Even when a targeting event resulted in homogenous reporter gene expression in presence of antibiotics imposing selective pressure, in most of these clones mosaic gene silencing started when this selective pressure was removed (20). These and other observations indicate that standard transfection procedures mostly result in integration events in chromosomal loci that do not warrant constant transgene expression, in particular without selection for a co-integrated resistance marker gene. Mosaic expression patterns are often a consequence of dynamic epigenetic settings proximal to the transgene. It is reasonable to assume that selection marker genes at these sites show similar expression patterns. Drug selection would result in a proliferative advantage for the subset of cells expressing an antibiotic resistance gene at a given time, rather than actively fending off the dominant influence of repressive chromatin structures. While such loci might still show a drug-selection enforceable, ostensible stability of gene expression, they are clearly not suited to support the analysis of transgenic *cis* elements contributing to mosaicism. With our approach, cell lines were isolated on the basis of persistent transgene expression in the first place, and not by antibiotic resistance. We assume that they would depend on favorable integration sites that are not exposed to such invasion of repressive chromatin modifications or potentially also DNA methylation.

To test this concept we generated stable cell lines using GFP as a sortable marker protein for FACS. Far from establishing itself as a routine procedure, the potential of this methodology for the generation of stable cell lines has been demonstrated, often also in the enrichment of subpopulations of cells showing particular high expression levels (24–30). Starting from a single transfection of a standard GFP reporter construct, we showed a direct comparison between a cell sorting- and an antibiotic selection-based approach for the generation of stable cell lines. GFP-positive cells in sorted cell population appeared at a much higher frequency as compared to those obtained by drug selection. This is the expected consequence of the isolation procedure, which is based directly on the expression of the fluorescent protein. However, even after clonal outgrowth, positive cell lines showed homogenous, stable GFP expression in absence of selective agents, unlike the cell lines obtained by standard drug selection procedures. Our interpretation is that the cell sorting and subcloning procedure described will specifically enrich for integration events in chromosomal loci, where transgenes are invisible to cellular mechanisms responsible for triggering or mediating silencing, allowing them to 'survive' repeated cell sorting. It will be interesting to identify these favorable chromosomal sites for transgene integration and determine their localization relative to functional entities within the genome (e.g. transcription units, matrix attachment regions, heterochromatic regions) and use well characterized loci with the desirable properties for retargeting by chromosome engineering (31,32). Our parallel experiments using a drug-selection protocol demonstrate that these integration events are

lost in the course of antibiotic selection. With its own transcription unit, the criteria for resistance marker expression obviously do not coincide with those for an independently expressed gene of interest, even when placed in close proximity. Threshold effects in mediating drug resistance as compared to the quantitative readout possible by cell sorting may also contribute to the apparent imbalance of integration sites recovered through the different protocols.

The isolation of stably transfected cells by FACS-based cell sorting for GFP-expressing cells and the resulting enrichment of transgenes integrated in a favorable chromatin environment is obviously also a form of selection. One could argue that given the readout used, this isolation procedure is simply a more stringent selection principle than the drug selection used by us. More stringent and quantitative drug-based approaches for cell line selection might therefore also improve the performance of stable transformants. Without having tested this directly, both the outcome of our experiments as well as the concept of designing any given selection scheme based on the gene of interest itself, already argue that the integration of sorting strategies in gene transfer protocols should be considered an attractive experimental option in general.

Whatever the exact requirements for persistent, uniform expression of transgenes might be, it is clear that such conditions are not necessarily precluded by some of the factors often suspected to contribute to mosaicism and transcriptional silencing. For example, bacterial vector sequences have been implicated to play a major role in these processes (33–36) and viral enhancer elements like that found in the CMV immediate early promoter are often deemed to be more susceptible to silencing than cellular expression signals like the EF promoter, both *in vitro* and *in vivo* (34,37–39). Such a trend was also seen in our experiments using FACS, where the use of the CMV promoter as well as inclusion of the plasmid backbone had a negative effect on the uniformity of gene expression. However, this effect was rather small and clearly did not prevent formation of homogeneously positive clones under these conditions.

Cell line generation with FACS or MACS is mostly restricted to transgenes amenable to the respective separation principle. This limitation can be overcome by the recombinase-mediated deletion of the sortable marker gene, reminiscent of stop cassettes mostly used to conditionally activate transgenes in mice, depending on a recombinase expression pattern (40). Related strategies have also been described for cell lines (41,42). Such approaches have the additional advantage that they facilitate the creation of stable genetic conditions without expression of the transgene of interest itself, circumventing possible negative effects on proliferation during cell line establishment. Our solution for stable expression of non-sortable genes is not limited to using strong constitutive promoters meant to facilitate cell sorting. The EF1 α promoter can also be placed together with the GFP marker between the flanking FRT sites. Upon deletion of this whole transcription unit, any upstream promoter (e.g. an inducible or cell type specific) could take over expression of the gene of interest.

Thus, the strategy for establishing stable cell lines presented here can be widely applied, resulting in clones displaying expression characteristics often exceedingly difficult to obtain by antibiotic-enforced selection protocols. The basic procedure of isolating cell clones (repeated FACS, subcloning, reporter deletion and recloning) can be completed in about 6 weeks. Depending on the questions asked or the problem confronted with, establishing such cell lines will often be a worthwhile investment of time. It will be interesting to further evaluate cell sorting as the basic principle for generating cell lines as an additional parameter in optimizing recombinant protein production (43). Other examples for the utility of this route of cell line production include the need to grow cells under antibiotic-free conditions, be it to minimize cellular stress (44) or to eliminate the risk of antibiotic contaminants in industrial fermentation processes. For cell-based assays, the issue of homogenous transgene expression can also have significant consequences. While heterogeneous transgene expression might be acceptable when transgene expression results in an activation of a cellular response, it is unacceptable if the expressed transgene is e.g. a repressor protein or a microRNA, downregulating particular cellular targets. Assuming expression exists in only half of the cells, the response as measured for the cell population will never exceed a factor of two. This will frequently be prohibitive for any kind of conclusive analysis and emphasizes the need to control the variability of transgene expression in cellular systems.

ACKNOWLEDGEMENTS

All members of our laboratory are acknowledged for stimulating discussions. We thank Jürgen Bode, HZI Braunschweig, for providing reagents and Chris Kaufman for critically reading the manuscript. This project was supported by grants from the VolkswagenStiftung (Conditional mutagenesis) and the BMBF (START-MS). Funding to pay the Open Access publication charges for this article was provided by internal funds of the Max Delbrück Center.

Conflict of interest statement. None declared.

REFERENCES

- Makrides, S.C. (2003) *Gene Transfer and Expression in Mammalian Cells*. Elsevier B.V., Amsterdam.
- Colosimo, A., Goncz, K.K., Holmes, A.R., Kunzelmann, K., Novelli, G., Malone, R.W., Bennett, M.J. and Gruenert, D.C. (2000) Transfer and expression of foreign genes in mammalian cells. *Biotechniques*, **29**, 314–318, 320–312, 324 passim.
- Whitelaw, E., Sutherland, H., Kearns, M., Morgan, H., Weaving, L. and Garrick, D. (2001) Epigenetic effects on transgene expression. *Methods Mol. Biol.*, **158**, 351–368.
- Martin, D.I. and Whitelaw, E. (1996) The vagaries of variegating transgenes. *Bioessays*, **18**, 919–923.
- Miranda, T.B. and Jones, P.A. (2007) DNA methylation: the nuts and bolts of repression. *J. Cell Physiol.*, **213**, 384–390.
- Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693–705.
- Artelt, P., Grannemann, R., Stocking, C., Friel, J., Bartsch, J. and Hauser, H. (1991) The prokaryotic neomycin-resistance-encoding gene acts as a transcriptional silencer in eukaryotic cells. *Gene*, **99**, 249–254.
- Henikoff, S. (1998) Conspiracy of silence among repeated transgenes. *Bioessays*, **20**, 532–535.
- Ill, C.R. and Chiou, H.C. (2005) Gene therapy progress and prospects: recent progress in transgene and RNAi expression cassettes. *Gene Ther.*, **12**, 795–802.
- Yew, N.S. (2005) Controlling the kinetics of transgene expression by plasmid design. *Adv. Drug Deliv. Rev.*, **57**, 769–780.
- Newlands, S., Levitt, L.K., Robinson, C.S., Karpf, A.B., Hodgson, V.R., Wade, R.P. and Hardeman, E.C. (1998) Transcription occurs in pulses in muscle fibers. *Genes Dev.*, **12**, 2748–2758.
- Nelson, D.E., Ihekwaba, A.E., Elliott, M., Johnson, J.R., Gibney, C.A., Foreman, B.E., Nelson, G., See, V., Horton, C.A., Spiller, D.G. *et al.* (2004) Oscillations in NF-kappaB signaling control the dynamics of gene expression. *Science*, **306**, 704–708.
- Golding, I. and Cox, E.C. (2006) Eukaryotic transcription: what does it mean for a gene to be 'on'? *Curr. Biol.*, **16**, R371–373.
- Kaern, M., Elston, T.C., Blake, W.J. and Collins, J.J. (2005) Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.*, **6**, 451–464.
- Fiering, S., Whitelaw, E. and Martin, D.I. (2000) To be or not to be active: the stochastic nature of enhancer action. *Bioessays*, **22**, 381–387.
- Strathdee, C.A., McLeod, M.R. and Underhill, T.M. (2000) Dominant positive and negative selection using luciferase, green fluorescent protein and beta-galactosidase reporter gene fusions. *Biotechniques*, **28**, 210–212, 214.
- Schlake, T. and Bode, J. (1994) Use of mutated FLP recognition target (FRT) sites for the exchange of expression cassettes at defined chromosomal loci. *Biochemistry*, **33**, 12746–12751.
- Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A. and Tsien, R.Y. (2002) A monomeric red fluorescent protein. *Proc. Natl Acad. Sci. USA*, **99**, 7877–7882.
- Buchholz, F., Angrand, P.O. and Stewart, A.F. (1998) Improved properties of FLP recombinase evolved by cycling mutagenesis. *Nat. Biotechnol.*, **16**, 657–662.
- Liu, W., Xiong, Y. and Gossen, M. (2006) Stability and homogeneity of transgene expression in isogenic cells. *J. Mol. Med.*, **84**, 57–64.
- Kirsch, P., Hafner, M., Zentgraf, H. and Schilling, L. (2003) Time course of fluorescence intensity and protein expression in HeLa cells stably transfected with hrGFP. *Mol. Cells*, **15**, 341–348.
- Gaines, P. and Wojchowski, D.M. (1999) pIRES-CD4t, a dicistronic expression vector for MACS- or FACS-based selection of transfected cells. *Biotechniques*, **26**, 683–688.
- Barnes, L.M., Bentley, C.M. and Dickson, A.J. (2003) Stability of protein production from recombinant mammalian cells. *Biotechnol. Bioeng.*, **81**, 631–639.
- Browne, S.M. and Al-Rubeai, M. (2007) Selection methods for high-producing mammalian cell lines. *Trends Biotechnol.*, **25**, 425–432.
- Van Tendeloo, V.F., Ponsaerts, P., Van Broeckhoven, C., Berneman, Z.N. and Van Bockstaele, D.R. (2000) Efficient generation of stably electrotransfected human hematopoietic cell lines without drug selection by consecutive FACSorting. *Cytometry*, **41**, 31–35.
- Gubin, A.N., Reddy, B., Njoroge, J.M. and Miller, J.L. (1997) Long-term, stable expression of green fluorescent protein in mammalian cells. *Biochem. Biophys. Res. Commun.*, **236**, 347–350.
- Migliaccio, A.R., Bengra, C., Ling, J., Pi, W., Li, C., Zeng, S., Keskinetepe, M., Whitney, B., Sanchez, M., Migliaccio, G. *et al.* (2000) Stable and unstable transgene integration sites in the human genome: extinction of the green fluorescent protein transgene in K562 cells. *Gene*, **256**, 197–214.
- Brezinsky, S.C., Chiang, G.G., Szilvasi, A., Mohan, S., Shapiro, R.I., MacLean, A., Sisk, W. and Thill, G. (2003) A simple method for enriching populations of transfected CHO cells for cells of higher specific productivity. *J. Immunol. Methods*, **277**, 141–155.
- Meng, Y.G., Liang, J., Wong, W.L. and Chisholm, V. (2000) Green fluorescent protein as a second selectable marker for selection of high producing clones from transfected CHO cells. *Gene*, **242**, 201–207.
- Mancia, F., Patel, S.D., Rajala, M.W., Scherer, P.E., Nemes, A., Schieren, I., Hendrickson, W.A. and Shapiro, L. (2004) Optimization of protein production in mammalian cells with a coexpressed fluorescent marker. *Structure*, **12**, 1355–1360.

31. Bouhassira, E.E., Westerman, K. and Leboulch, P. (1997) Transcriptional behavior of LCR enhancer elements integrated at the same chromosomal locus by recombinase-mediated cassette exchange. *Blood*, **90**, 3332–3344.
32. Wirth, D., Gama-Norton, L., Riemer, P., Sandhu, U., Schucht, R. and Hauser, H. (2007) Road to precision: recombinase-based targeting technologies for genome engineering. *Curr. Opin. Biotechnol.*, **18**, 411–419.
33. Townes, T.M., Lingrel, J.B., Chen, H.Y., Brinster, R.L. and Palmiter, R.D. (1985) Erythroid-specific expression of human beta-globin genes in transgenic mice. *Embo J.*, **4**, 1715–1723.
34. Teschendorf, C., Warrington, K.H. Jr, Siemann, D.W. and Muzyczka, N. (2002) Comparison of the EF-1 alpha and the CMV promoter for engineering stable tumor cell lines using recombinant adeno-associated virus. *Anticancer Res.*, **22**, 3325–3330.
35. Chen, Z.Y., He, C.Y., Meuse, L. and Kay, M.A. (2004) Silencing of episomal transgene expression by plasmid bacterial DNA elements in vivo. *Gene Ther.*, **11**, 856–864.
36. Chen, Z.-Y., Riu, E., He, C.-Y., Xu, H. and Kay, M.A. (2008) Silencing of episomal transgene expression in liver by plasmid bacterial backbone DNA is independent of CpG methylation. *Mol. Ther.*, **16**, 548.
37. Gopalkrishnan, R.V., Christiansen, K.A., Goldstein, N.I., DePinho, R.A. and Fisher, P.B. (1999) Use of the human EF-1alpha promoter for expression can significantly increase success in establishing stable cell lines with consistent expression: a study using the tetracycline-inducible system in human cancer cells. *Nucleic Acids Res.*, **27**, 4775–4782.
38. Gill, D.R., Smyth, S.E., Goddard, C.A., Pringle, I.A., Higgins, C.F., Colledge, W.H. and Hyde, S.C. (2001) Increased persistence of lung gene expression using plasmids containing the ubiquitin C or elongation factor 1alpha promoter. *Gene Ther.*, **8**, 1539–1546.
39. Byun, H.M., Suh, D., Jeong, Y., Wee, H.S., Kim, J.M., Kim, W.K., Ko, J.J., Kim, J.S., Lee, Y.B. and Oh, Y.K. (2005) Plasmid vectors harboring cellular promoters can induce prolonged gene expression in hematopoietic and mesenchymal progenitor cells. *Biochem. Biophys. Res. Commun.*, **332**, 518–523.
40. Lakso, M., Sauer, B., Mosinger, B. Jr, Lee, E.J., Manning, R.W., Yu, S.H., Mulder, K.L. and Westphal, H. (1992) Targeted oncogene activation by site-specific recombination in transgenic mice. *Proc. Natl Acad. Sci. USA*, **89**, 6232–6236.
41. Spiotto, M.T. and Schreiber, H. (2006) Floxed reporter genes: Flow-cytometric selection of clonable cells expressing high levels of a target gene after tamoxifen-regulated Cre-loxP recombination. *J. Immunol. Methods.*, **312**, 201–208.
42. Brough, R., Papanastasiou, A.M. and Porter, A.C. (2007) Stringent and reproducible tetracycline-regulated transgene expression by site-specific insertion at chromosomal loci with pre-characterised induction characteristics. *BMC Mol. Biol.*, **8**, 30.
43. Wurm, F.M. (2004) Production of recombinant protein therapeutics in cultivated mammalian cells. *Nat. Biotechnol.*, **22**, 1393–1398.
44. Rodolose, A., Barbat, A., Chantret, I., Lacasa, M., Brot-Laroche, E., Zweibaum, A. and Rousset, M. (1998) Selecting agent hygromycin B alters expression of glucose-regulated genes in transfected Caco-2 cells. *Am. J. Physiol.*, **274**, G931–938.