Chromatin architecture and the orchestration of gene expression: cell systems to explore epigenetic gene control

Brink, M.C.

Publication date
2009

Document Version
Final published version

Chapter 5

A cell system to analyze epigenetic gene regulation

Maartje C. Brink, Diewertje G. E. Piebes, Hinco J. Gierman, Mireille H. Indemans, Angelique Meulemans-Rops, Giuseppina Giglia-Mari, Rogier Versteeg, Roel van Driel, Johan van der Vlag, Wim Vermeulen & Pernette J. Verschure

Adapted from patent application PCT/NL2008/050150

Abstract

Packing of the eukaryotic genome into higher-order chromatin structures is tightly related to gene expression, although the underlying molecular mechanisms are currently unclear. The lac-operator/lac-repressor system allows visualization of a specific chromatin domain as well as manipulation of such a domain by the targeting of epigenetic regulatory proteins. To systematically study the causal relationship between changes in the epigenetic state, chromatin structure and gene expression within a genomic environment, we engineered and integrated a composite targeting cassette in human cells. The cassette contains a reporter gene and two short arrays of protein-binding sites that enable targeting of fluorescent epigenetic regulatory proteins and allow manipulation of the epigenetic state and visualization of the cassette. Criteria for the functionality of the cell system are (i) single integration of the cassette at a designated site, (ii) visualization of the integrated cassette, (iii) ability to measure changes in chromatin structure and (iv) ability to measure transcriptional activity. As a proof of principle, two mono-clonal cell lines were established with a single integration using Flp recombinase in genomic environments of either high or of low gene activity. The use of fluorescent *in situ* hybridization and reverse transcription quantitative PCR enable the measurement of changes in chromatin structure and changes in gene activity, respectively. Systematic and quantitative *in vivo* measurements using this system will provide insight into causal relationships between an induced epigenetic state, changes in chromatin structure and transcriptional activity in a specific genomic environment.

Introduction

Chromatin organization plays an important role in mammalian gene regulation, although we lack clear understanding of cause-and-effect relationships (for a review, see Goetze *et al.*, 2007b). Studies indicate that silent, gene-poor chromatin is positioned more to the periphery of the nucleus, whereas active, gene-rich chromatin adopts a more central position (Finlan *et al.*, 2008; Goetze *et al.*, 2007a; Guelen *et al.*, 2008). Furthermore, chromatin condensation was seen to correlate with an epigenetic state marked by H3K9 methylation and gene repression (Verschure *et al.*, 2005). *Vice versa*, chromatin decondensation has been associated with transcriptional activity and specific H3 acetylation marks (Tumbar *et al.*, 1999). In addition, looping of genes out of their chromosome territory has been observed for several transcriptionally active genes, such as the functionally related genes of the HoxB cluster that can loop out sequentially upon induction, facilitating transcriptional co-regulation (Chambeyron and Bickmore, 2004). At the linear DNA level, clustering of functionally unrelated genes was

identified (Caron *et al.*, 2001). Specific chromosomal regions containing clusters of highly expressed gene of unrelated function were named ridges (regions of increased gene expression), and clusters of low-expressed genes anti-ridges. Integration of a GFP reporter gene in a ridge region resulted in an average four-fold higher expression of the reporter compared to expression when integrated in an anti-ridge, suggesting that the genomic environment contributes considerably to the transcriptional potential of a gene (Gierman *et al.*, 2007).

The advantage of the use of the lac operator/lac repressor system is that it provides information on the causal relationships between chromatin organization and gene activity (Robinett *et al.*, 1996). Initially, Belmont and coworkers created an artificial chromosomal domain in mammalian cells consisting of an integrated and amplified array of bacterial lactose operator (lacO) sites, flanked by a dihydrofolate (DHFR) selection marker. Binding of the lac repressor (lacR)-GFP fusion protein to the amplified chromosomal domain allowed the *in vivo* visualization of large-scale chromatin structure in mammalian cells (Robinett *et al.*, 1996). The lacO/lacR targeting system has been used in a number of studies illuminating the interplay between the higher-order packaging of chromatin and gene activity (Brink *et al.*, 2006; Tumbar *et al.*, 1999; Verschure *et al.*, 2005; and chapter 2 of this thesis). Furthermore, the targeting system has also been used to analyze the relationship between gene activity and positioning of genes in the nucleus. Tumbar and Belmont (2001) showed that after gene activation, the array was repositioned to a more central location in the nucleus. Other studies showed that targeting of the integrated lacO system to the nuclear lamina caused repression of a reporter gene and of some, but not all, of the genes near the targeting site (Finlan *et al.*, 2008; Reddy *et al.*, 2008). In a related study, it was shown that reporter gene transcription could be induced at the nuclear lamina, indicating that nuclear localization is important, but not decisive for gene regulation (Kumaran and Spector, 2008).

Although our understanding of epigenetic regulatory mechanisms has expanded considerably, we still lack insight into the effect of the local genomic environment on gene regulation. In this study we created a cell system consisting of an integrated cassette in human cells to systematically measure the causal relationship between changes in the epigenetic state, chromatin structure and gene expression in a specific genomic environment. The cassette contains a reporter gene and two short arrays of non-endogenous protein-binding DNA sites (the bacterial lactose operator (lacO) sites and the yeast Upstream Activating Sites (UAS)) that enable targeting of fluorescent epigenetic regulatory proteins, allowing induction of a specific epigenetic state and the visualization of the cassette. Criteria for the functionality of the cell system are (i) single integration of the cassette at a designated site, (ii) visualization of the integrated cassette in living cells, (iii) ability to measure changes in chromatin structure and (iv) ability to measure transcriptional activity. As a proof of principle, two mono-clonal cell lines were established with a single integration using Flp recombinase in specific

genomic sites of either high or low gene activity, *i.e.* a ridge and an anti-ridge, respectively. (Gierman *et al.*, 2007; Gronostajski and Sadowski, 1985; Sauer, 1994). The clones contain the integrated cassette in functionally different genomic regions to assess the contribution of the local chromatin environment on epigenetic gene control. We demonstrate the visualization of the lacO repeat by *in vivo* detection upon lacR-GFP binding and by fluorescent *in situ* hybridization (FISH). We provide data on gene activity of the reporter gene in the established cell lines by means of reverse transcription quantitative PCR (RT-qPCR) Our data show that when the cassette is integrated in a ridge region, reporter gene activity is two-fold higher than when integrated in an anti-ridge. We conclude that the criteria for useful functioning of the cell system are fulfilled and we discuss suggestions for additional technical improvements of the cell system. This system will be an important tool for gaining insight into the molecular mechanisms underlying epigenetic gene regulation in a controlled and systematic manner.

Materials & Methods

Cloning of plasmids

Plasmid pPS-8.4, a pUC18 derivative containing 4 copies of the lacO octamer divided by DNA spacer fragments, was used as a cloning backbone vector (Robinett *et al.*, 1996). A dsRed-MTS-IRES-hygromycin reporter cassette was constructed by inserting a pDsRed2-Mito-fragment (Clontech, USA) with *SnaBI* and filled-in *NotI* ends into *SnaBI-SmaI* sites of pIREShyg3 (Clontech, USA). The resulting reporter cassette was excised with *SalI* and inserted into the pPS-8.4 *XhoI* site. Next, the Flp Recombination Target (FRT) site (Gronostajski and Sadowski, 1985; Sauer, 1994) was synthesized with *ClaI-NruI* restriction sites 5' and *PacI-XmaI* sites 3' and subcloned into *ClaI-XmaI* of pBluescript II SK+ (Stratagene). *KpnI-MluI-SwaI* sites were added to the 5' end and *ClaI* was added to the 3' end of a 14x repeat of the UAS operator by PCR on the pZG18 plasmid (obtained from Carl Schildkraut, Department of Cell Biology, Albert Einstein college of Medicine, New York, USA). The UAS repeat was ligated into *KpnI-Cla*I of FRT-containing pBluescript. UASFRT was excised with *MluI-XmaI* and ligated into *MluI-XmaI* of p8.4dsRed, creating p32RUF (figure 1). All culturing steps involving the lacO repeat were performed at 30°C using MAX efficiency Stbl2 competent cells (Invitrogen). All products obtained by PCR or oligo synthesis were sequenced to ensure fidelity.

Cell culture, Lox and Flip-In

Cell lines H100 and H94 are human embryonal kidney (HEK293) cells containing a lentiviral construct, integrated in a ridge (HG17: position 147.192.859 in forward orientation) or anti-ridge (HG17: position 218.918.221 in reverse orientation) region, respectively (Gierman *et al.*, 2007). The lentiviral construct contained an FRT site (Flp Recombinase Target site; O'Gorman *et al.*, 1991) to promote homologous recombination at the site of integration, and a GFP reporter cassette flanked by two loxP sites. In order to insert p32RUF, the GFP reporter cassette was loxed out by transfection with Cre recombinase. Cells that had lost the GFP cassette were identified by fluorescence activated cell sorting and single cells were seeded in a 96-well plate. Next, p32RUF was inserted at the FRT integration site by Flp-In as described previously (O'Gorman *et al.*, 1991). Briefly, p32RUF was cotransfected with Flp recombinase-expressing plasmid pOG44 (Invitrogen), required for site-specific recombination at the FRT site, in a 9:1 ratio (pOG44:p32RUF) using Lipofectamine 2000 (Invitrogen) according to the manufacturer's protocol. Cells were diluted 1:10 after 24 hours and selection pressure (115 μg/ml hygromycin) was added. After 2 weeks, individual clones could be distinguished. Clones were transferred to a 12-well for further culturing and DNA isolation. Cell lines were tested for integration at the FRT site by PCR and southern blotting. VH10 and HEK cells were

cultured in Dulbecco Minimal Essential Medium (DMEM) containing 10% or 15% fetal bovine serum respectively, 1% penicillin/streptomycin and 1% glutamin at 37°C in a 10% CO_2 atmosphere. For RT-qPCR and FISH, HEK cells were transfected at 80% confluency with Lipofectamine 2000 as described by the manufacturer. After 24-48 hours, cells were either harvested for RNA isolation or transferred to poly-L-lysine coated slides in preparation for FISH.

Southern Blot
Isolation of genomic DNA was performed by overnight cell lysis at 55°C in Tris buffer (10 mM Tris pH 7.5, 2 mM EDTA, 10 mM NaCl, 0.5% SDS) supplemented with 1 mg/ml Proteinase K (Roche) and subsequent precipitation. Genomic DNA was cut overnight with *BglII* (Roche, 50 U/µl) or a combination of *ClaI* and *XhoI* (Roche, 10 and 40 U/µl, respectively), so as to cut once in the lentiviral LTR and once in the integrated p32RUF, excising a single ~5.9 and ~4 kb band, respectively. 10 µg of restricted DNA was run overnight on a 1% agarose gel and transferred to an Amersham Nylon membrane the following night. The blot was probed overnight with a ~300 bp lacO fragment, labeled with $AT^{32}P$. Washes were performed at 63°C and detection occurred with the aid of a phosphor image screen.

Reverse transcription quantitative PCR (RT-qPCR)
RNA was isolated after trypsinization from 25 cm^2 culture flasks with the Qiagen RNeasy Mini Kit according to the accompanying manual. After solution in RNAse-free H_2O, RNA was processed immediately for cDNA synthesis: 5 µg of RNA was incubated for 20 minutes with RNAse-free DNAse I (Roche) at room temperature. Next, we heated the RNA for 5 minutes at 70°C to dissociate the tertiary structure followed by immediate cooling on ice for 5 minutes. cDNA was synthesized with the Improm II kit (Promega), using oligo dT primer and incubating the reaction mix for 5 minutes at 25°C to anneal primers and 60 minutes at 42°C to allow cDNA synthesis. The reaction was inactivated for 15 minutes at 70°C and cooled on ice. Quantitative PCR reactions were performed on a 7500 Real Time Cycler (Applied Biosystems), using accompanying software and SYBR-green (Invitrogen) as a detection reagent. Relative gene expressions were calculated using the ΔΔCT method.

Three-dimensional fluorescent in situ hybridization (3D FISH)
For 3D FISH, cells were grown on poly-L-lysine coated slides and fixed in 4% paraformaldehyde for 10 minutes at 4°C. Paraformaldehyde was quenched after 3 PBS washes in 100 mM glycin for 10 minutes and cells were subsequently placed in 0.1 N HCl, washed in PBS and permeabilized in PBS containing 0.5% Triton X-100 and 0.5% Saponin for 20 minutes. After three PBS washes, slides were treated with 100 µl RNAse (100 µg/ml) at 37°C for 30 minutes. Three PBS washing steps were followed by a 5-minute incubation in 2x SSC for 5 minutes and immersion in 50% formamide in 2x SSC for up to 2 hours. For preparation of the probe 100 ng of BAC or 10 ng of lac DNA, amplified by DOP-PCR and labeled by nick translation with either biotin of digoxigenin, was dissolved in deionised formamide supplemented with 50% formamide and 10% dextrane sulfate to a final concentration of 12.5 ng probe/µl. The probe was denatured at 80°C for 10 minutes and placed on ice, followed by 1 to 2 hours of pre-annealing at 37°C. Next, the probe was added to the pre-treated slide and codenatured for 2 minutes at 80°C. Hybridization was allowed to proceed overnight in a moist chamber at 37°C. Three posthybridization washes were carried out with 2x SSC-50% formamide at 45°C, followed by three washes in 0.1x SSC at 60°C and blocking with 4x SSC containing 5% non-fat dried milk for 1 hour at room temperature. All incubations for probe detection were performed for 20 minutes at room temperature in 4x SSC containing 5% (wt/vol) non-fat dried milk, alternated by three washes in 0.05% Tween/2x SSC. Slides were finally labeled in DAPI (4', 6'-diamidino-2-phenylindole; Vector Laboratories, Burlingame, CA) and embedded in Vectashield (Brunschwig, Burlingame, CA). Slides were imaged using a Zeiss LSM 510 (Zeiss, Jena, Germany) confocal laser scanning microscope equipped with a 63x/1.4 oil immersion objective. We used multitrack scanning, employing a UV laser (364 nm), an argon laser (488 nm) and a helium-neon laser (543 nm) to excite DAPI staining and green and red fluorochromes. Emitted fluorescence was detected with BP 385-470, BP 505-550 and 560 LP filters. Three-dimensional images were scanned at 512 by 512 using 400 nm axial and 60 nm lateral sampling rates. Images were averaged 4 times.

Results and discussion

Development and single integration of the cell system

To study the *in vivo* causal relationships between local induction of an epigenetic chromatin state and the changes in chromatin structure and gene activity in a specific genomic context, we created a cell system consisting of a cassette containing two short arrays of non-endogenous protein-binding sites and a reporter gene. The cassette (p32RUF) contains 32 tandem copies of the lacO sequence (lacO repeat) and 14 tandem copies of the UAS site for targeting by lac repressor (lacR) and Gal4, respectively. In addition, the targeting cassette contains a dsRed reporter gene tagged by a mitochondrial targeting signal (MTS) to reduce fluorescent background in the nucleus, an FRT site to enable homologous recombination at defined genomic sites and a hygromycin selection marker for subsequent selection (figure 1). To test the functionality regarding visualization, ability to measure chromatin folding and gene activity, the cassette was integrated in human embryonal kidney 293 (HEK 293) cells. Integration of foreign DNA into defined genomic regions of human cells is very inefficient. We used HEK cell lines containing an FRT site promoting recombination at a defined position on chromosome 1 band q21.1 and chromosome 2 band q36.1, corresponding to a ridge and anti-ridge (Gierman *et al.*, 2007). These cell lines additionally contained a GFP reporter used for a previous study between two LoxP sites, which we removed using Cre/loxP recombinase. Next, we integrated the cassette into the FRT site with the aid of Flpase. The mono-clonally generated cell lines were screened for single integration by PCR and southern blotting. The HEK 94R99 and 100R21 clones were identified as clones having a single integration in the anti-ridge and in the ridge genomic region, respectively (figure 2).

During the multiple rounds of transfection (*i.e.* integration of an FRT and GFP reporter gene between LoxP sites, subsequent removal of the GFP reporter, integration of the composite cassette at the FRT site and the transfection with GFP-tagged lacR) the cells became increasingly difficult to culture and transfect, limiting the use of the system. The use of cell lines that enable homologous recombination, such as mouse embryonic stem cells, might be a suitable alternative to solve this issue. Moreover, the stable expression of inducible lacR targeting fusion proteins will avoid transfecting cell lines for each new experiment, while the inducibility limits continuous over-expression of lacR-tagged epigenetic regulators.

Visualization of 32 lacO sites *in vivo*

The amount of tandem copies of the non-endogenous protein-binding sites was minimized to 32 lacO sequences as the integration of large exogenous stretches of repetitive DNA may interfere with epigenetic gene regulation. To visualize the lacO repeat in the HEK cell nuclei, images were taken from nuclei with a low EGFP-lacR expression level to minimize background fluorescence (figure 3). The lacO repeat is

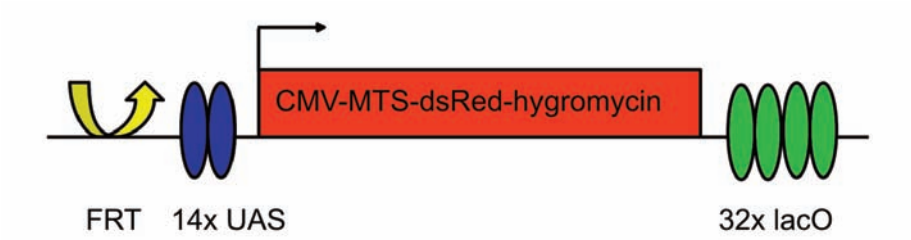

Figure 1. Schematic representation of the cassette p32RUF. The cell system consists of an integrated cassette containing a CMV-driven red fluorescent reporter with a mitochondrial targeting signal and a hygromycin selection marker separated by an IRES (internal ribosome entry site) to ensure translation initiation (red), flanked by 14 tandem copies of the non-endogenous UAS protein-binding sites (blue) upstream and 32 tandem copies of the lacO protein-binding sites (green) downstream. A Flp-recombinase targeting (FRT; yellow) site enables homologous recombination.

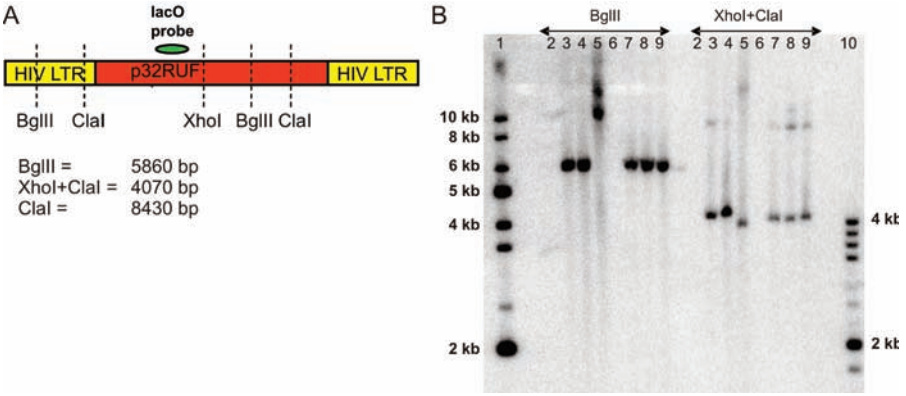

Figure 2. Southern blot analysis to verify the single integration of the p32RUF cassette. (A) Schematic representation of the position of the lacO probe (green) relative to restriction sites in the p32RUF cassette (red) that has been flipped into the virally integrated FRT, which is flanked by viral long terminal repeats (LTR; yellow). Expected fragment lengths for *BglII*, *XhoI*+*ClaI* and *ClaI* are given. (B) Genomic DNA of 6 cell lines with potential integration of p32RUF in an anti-ridge (lanes 3-5) or ridge (lanes 7-9) was digested overnight with *BglII* (first half of blot) or a combination of *XhoI* and *ClaI* (second half of blot). DNA fragments were separated by gel electrophoresis, transferred to a nylon membrane and probed for the lacO with an AT^{32}P-labeled probe. Lanes numbered 1 and 10 contain a DNA marker, lanes 2 and 6 contain a control cell line without lacO integration. The ~8.4 kb bands in the *XhoI/ClaI*-digested lanes are due to incomplete *XhoI* digestion (data not shown). Cell lines HEK 94R99 and 100R21, chosen for further experiments, were loaded in lanes 3 and 7, respectively.

indicated by the arrow.

In vivo visualization of the 32 tandem lacO copies integrated in the HEK cell lines proved to be challenging, as opposed to the clear detectability of EGFP-lacR binding at lacO arrays spanning several mega basepairs (Belmont *et al.*, 1999; Robinett *et al.*, 1996; Tumbar *et al.*, 1999). The lacR has a high affinity for the bacterial lacO sequence with a K_D between 10^{-11} and 10^{-13} M (Chen and Matthews, 1992; Riggs *et al.*, 1970). However, the lacR also binds non-specifically to DNA with dissociation constants

differing an estimated 6-8 orders of magnitude compared to specific binding (Lin and Riggs, 1975). The occupancy of a binding site depends on the affinity of the lacR for that site, the concentration of lacR molecules and on the number of binding sites. Given the relatively high amount of non-specific compared to specific binding sites, the absolute number of lacR molecules bound to non-specific DNA will be very high, thereby creating considerable background noise. We conclude that under optimal imaging conditions 32 tandem copies of the lacO site can be visualized in HEK cells if the concentration of the lacR-GFP protein is sufficiently low. To increase visibility, a higher amount of lacO sites may be used. Previously, it was shown that, in principle, a minimal amount of a 128 lacO sequences is detectable in yeast and mammalian cells (Brickner and Walter, 2004; Chubb *et al.*, 2002). Alternatively, the concentration of lacR molecules can be significantly lowered, such that only several tens of lacR molecules, which will mainly occupy lacO sites, are available in each cell (Elf *et al.*, 2007).

Chromatin structure visualized by FISH

Measurements on the distance between two positions on a single chromosome provide information on how chromatin is folded locally. Changes in the distance between these two positions after targeting epigenetic regulatory proteins are indicative of changes in chromatin folding due to the actions of these regulators. We therefore set out to detect the integrated cassette and a genomic reference site at a defined distance from the integrated region on the same chromosome by 3D FISH. We used a plasmid probe to detect the cassette and a specific BAC probe to recognize the reference site. Figure 4 illustrates the visualization of a ~100 kb BAC probe and a ~9 kb probe of the integrated plasmid. Results indicate that the 3D FISH method allows distance measurements between the lacO repeat and a genomic region nearby, recognized by a BAC probe, generating information on chromatin compaction (Mateos-Langerak *et al.*, 2009, PNAS, in press). Ultimately, measuring distances between the loci labeled by lacR and Gal4 may generate information about the *in vivo* compaction of the chromatin fiber without using the FISH method. However, the non-endogenous protein-binding sites in our current pilot system are only ~4 kb apart at linear distance, which is not sufficient to measure physical distances using standard light microscopy techniques.

The transcriptional activity of a dsRed reporter gene is influenced by its genomic context

To quantify gene activity, we have measured the fluorescence intensity of the mitochondrial dsRed reporter by fluorescence activated cell sorting (FACS) analysis. For this purpose, a mouse fibroblast cell line (VH10) was established containing multiple random, stable integrations of the lacO repeat next to the dsRed-MTS reporter. We transfected the cells with EGFP-lacR-VP16, or EGFP-lacR as a control, and

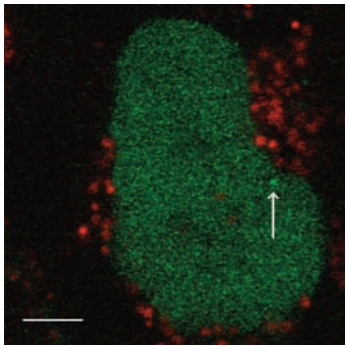

Figure 3. Living cell analysis of the integrated cassette. HEK cells with the p32RUF cassette flipped into the FRT site situated in a ridge, were transfected with EGFP-lacR. DsRed is expressed in the mitochondria. 3D images were recorded of nuclei with low EGFP-lacR expression with the pinhole set at 2.00 airy units to allow a maximal number of fluorescent proteins to be detected and averaged 8 times to reduce Poisson noise. The green signal represents EGFP-lacR. The arrow indicates the position of the lacO repeats targeted by EGFP-lacR. The image shown represents an individual optical section. Bar represents 2 µm.

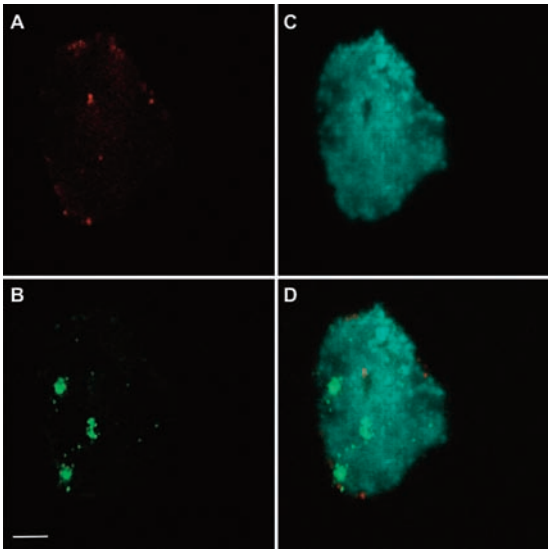

Figure 4. Fluorescent *in situ* hybridization. HEK cells with the p32RUF cassette flipped into the FRT site situated in an anti-ridge on chromosome 2 were FISH labeled. Probes containing p32RUF or nick-translated DNA from BAC RP11-465F2 (chromosome 11, q-arm), were labeled with biotin (A) or digoxigenin (B) respectively, and were hybridized to fixed and permeabilized cells. Antibody detection was achieved through Strep-Cy3 (Jackson, 1:1000) enhanced by goat anti-Strep-bio (Vector, 1:250) and Strep-Cy3 for biotin (A) and mouse anti-Dig (Boehringer, 1:250) followed by rabbit anti-mouse-FITC (Sigma, 1:1000) and goat anti-rabbit-Alexa488 (Molecular Probes, 1:200) for digoxigenin (B). 3D images were recorded in separate channels. The red signal represents the FISH labeled p32RUF probe recognized by anti-biotin antibodies, the green signal represents the FISH-labeled BAC probe recognized by anti-Dig antibodies and the blue signal represents DAPI staining. Part (D) portrays the merged image. Bar represents 2 µm.

A

epigenetic regulator	mean intensity of fluorescence	normalized intensity of fluorescence
EGFPlacR	7025	1.0
VP16	13775	2.0
HP1	6308	0.9
MeCP2	5249	0.8

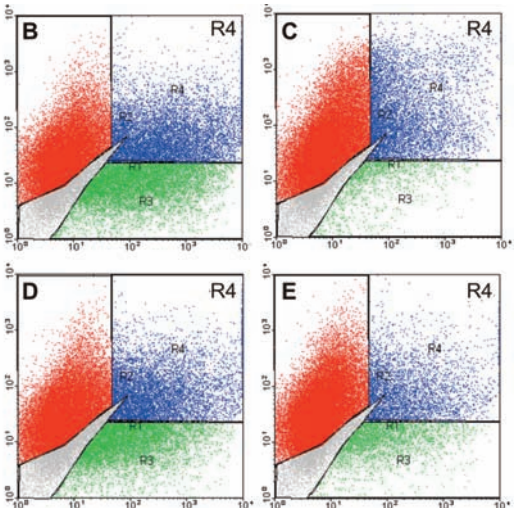

Figure 5. FACS analysis of dsRed reporter-expressing cells upon targeting. A VH10 cell line containing random integrations of 32 lacO repeats directly downstream of a dsRed reporter cassette was transiently transfected with EGFP-lacR control and EGFP-lacR-tagged VP16, HP1β and MeCP2. Fluorescence intensity was measured by FACS analysis; threshold values were set according to measurements in cells devoid of fluorescent proteins or expressing only dsRed or EGFP. (A) Table represents dsRed fluorescence intensities in EGFP-transfected cells. (B, C, D, E) Scatter plots of cells transfected with lacR (B), lacR-VP16 (C), lacR-HP1β (D) and lacR-MeCP2 (E) respectively. Green fluorescence is depicted along the *x*-axis and red fluorescence along the *y*-axis. Fluorescence intensities were measured in the R4-gated cells (blue).

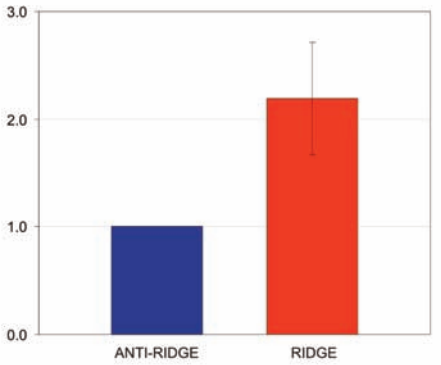

Figure 6. Relative reporter gene expression in the ridge versus anti-ridge region by RT-qPCR. RNA isolated from HEK 94R99 and 100R21 was reverse transcribed into cDNA with oligo dT primers. Relative dsRed expression from 4 different samples was determined by qPCR in a total of 23 experiments. Samples were normalized to RP L13A expression. DsRed expression in the anti-ridge region was set to 1.

analyzed dsRed expression of EGFP-expressing cells, 60 hours after transfection. The mean intensity of the dsRed fluorescence increased considerably in the VP16-expressing cells, going up to 13775 fluorescence units compared to 7205 units in the EGFP-lacR targeted control (figure 5). Next, we measured gene repression of the dsRed reporter by transiently transfecting the clones with EGFP-lacR-HP1β or EGFP-lacR-MeCP2. However, we were unable to detect an appreciable decrease in the expression levels of the reporter gene, measuring 6308 units for the HP1β-targeted and 5249 units for the MeCP2-targeted cells (EGFP-lacR = 7205 units, figure 5). The high expression and slow turnover of the dsRed protein (half life of ~5 days, Verkhusha *et al.*, 2003) prevents faithful detection of a decrease in the transcriptional activity of

the reporter gene. Therefore, we propose the use of an unstable reporter protein for future cell-system developments, for instance by fusion to a degradation signal (Lindsten *et al.*, 2003). Taken together, we conclude that in the current system fluorescence intensity of the dsRed reporter gene is an efficient read-out for gene activation, but not for gene repression.

To assess the influence of the local chromatin environment on gene expression, we measured RNA levels of the CMV-driven dsRed reporter after integration into HEK cell lines in two distinct genomic loci: a high and low expressing region, *i.e.* a ridge and anti-ridge, respectively. Reverse transcription quantitative PCR (RT-qPCR) was performed on HEK clones 94R99 and 100R21 for the dsRed gene, normalizing expression levels to ribosomal protein L13A (RPL13A; Pombo-Suarez *et al.*, 2008). We observed an approximately two-fold higher gene expression level of the dsRed protein when integrated in a ridge (chromosome 1 band q21.1) compared to an anti-ridge (chromosome 2 band q36.1; figure 6). These findings correspond to published data showing that expression levels of a GFP reporter gene driven by the PGK promoter are similarly affected by their integration into ridge or anti-ridge domains (Gierman *et al.*, 2007). It is interesting to note that the chromatin integration site of genes with a strong viral promoter (*i.e.* CMV) influences their expression to almost the same extent as a gene driven from the relatively weak PGK promoter (~three-fold difference when integrated in identical locations; H. Gierman, personal communication). These data support a model in which the transcriptional potential of a gene is partially determined by the chromatin environment in which it is embedded.

A drawback of determining RNA levels by biochemical bulk methods, such as RT-qPCR and northern blot analysis, is that they do not allow the detection of cell-to-cell variation. As an alternative, a recent advance has been made with the use of MS2 repeats, which form an RNA hairpin that is recognized by the MS2 coat protein, enabling single cell *in vivo* quantification of the expression level of a reporter gene by binding the fluorescently-tagged MS2 protein (Chubb *et al.*, 2006; Finlan *et al.*, 2008; Janicki *et al.*, 2004).

In conclusion, we have created a novel chromatin manipulation system in human cells consisting of a cassette that is integrated in a defined genomic background. We show preliminary data on the engineered cassette, integrated in single copy in a ridge and an anti-ridge region in human cell lines. We are able to (i) visualize the integrated cassette, (ii) determine the 3D position of the integrated cassette in relation to a nearby genomic reference region enabling measurement of the local chromatin compaction, (iii) measure changes in reporter gene activity by FACS analysis (gene activation) and RT-qPCR (gene activation and repression). We conclude that this cell system can be an important tool for studying molecular mechanisms of epigenetic gene regulation in living cells.

Acknowledgements

We would like to thank dr. M. J. Moné (VU, Amsterdam, The Netherlands) for expert assistance on the *in vivo* confocal imaging, ir. B. Hooijbrink (AMC, Amsterdam, The Netherlands) for performance of the FACS analysis, dr. M. S. Luijsterburg for critical reading of the manuscript and Ing. R. Bader and R. Charaffedine for technical assistance. This work was supported by the Netherlands Organisation for Scientific Research (NWO; project number VIDI2003/03921/ALW/016.041.311).