



# THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### Protecting a transgene expression from the HAC-based vector by different chromatin insulators

**Citation for published version:**

Lee, NC, Kononenko, AV, Lee, H-S, Larionov, V, Kouprina, N, Tolkunova, EN, Liskovykh, MA, Tomilin, AN, Masumoto, H & Earnshaw, WC 2013, 'Protecting a transgene expression from the HAC-based vector by different chromatin insulators' Cellular and Molecular Life Sciences., 10.1007/s00018-013-1362-9

**Digital Object Identifier (DOI):**

[10.1007/s00018-013-1362-9](https://doi.org/10.1007/s00018-013-1362-9)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Document Version:**

Publisher final version (usually the publisher pdf)

**Published In:**

Cellular and Molecular Life Sciences

**Publisher Rights Statement:**

© The Author(s) 2013. This article is published with open access at Springerlink.com

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Protecting a transgene expression from the HAC-based vector by different chromatin insulators

Nicholas CO Lee · Artem V. Kononenko · Hee-Sheung Lee · Elena N. Tolkunova · Mikhail A. Liskovykh · Hiroshi Masumoto · William C. Earnshaw · Alexey N. Tomilin · Vladimir Larionov · Natalay Kouprina

Received: 11 March 2013 / Revised: 25 April 2013 / Accepted: 30 April 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** Human artificial chromosomes (HACs) are vectors that offer advantages of capacity and stability for gene delivery and expression. Several studies have even demonstrated their use for gene complementation in gene-deficient recipient cell lines and animal transgenesis. Recently, we constructed an advance HAC-based vector, alphoid<sup>tet<sup>O</sup></sup>-HAC, with a conditional centromere. In this HAC, a gene-loading site was inserted into a centromeratin domain critical for kinetochore assembly and maintenance. While by definition this domain is permissive for transcription, there have been no long-term studies on transgene expression within centromeratin. In this study, we compared the effects of three chromatin insulators, cHS4,

gamma-satellite DNA, and tDNA, on the expression of an *EGFP* transgene inserted into the alphoid<sup>tet<sup>O</sup></sup>-HAC vector. Insulator function was essential for stable expression of the transgene in centromeratin. In two analyzed host cell lines, a tDNA insulator composed of two functional copies of *tRNA* genes showed the highest barrier activity. We infer that proximity to centromeratin does not protect genes lacking chromatin insulators from epigenetic silencing. Barrier elements that prevent gene silencing in centromeratin would thus help to optimize transgenesis using HAC vectors.

**Keywords** Insulator · tDNA-gamma-satellite · cHS4 · Human artificial chromosome-HAC

**Electronic supplementary material** The online version of this article (doi:10.1007/s00018-013-1362-9) contains supplementary material, which is available to authorized users.

N. C. Lee · A. V. Kononenko · H.-S. Lee · V. Larionov (✉) · N. Kouprina (✉)  
Laboratories of Molecular Pharmacology, National Cancer Institute, Bethesda, MD 20892, USA  
e-mail: larionov@mail.nih.gov

N. Kouprina  
e-mail: kouprinn@mail.nih.gov

E. N. Tolkunova · M. A. Liskovykh · A. N. Tomilin  
Institute of Cytology, Russian Academy of Sciences, St. Petersburg, Russia

H. Masumoto  
Laboratory of Cell Engineering, Department of Human Genome Research, Kazusa DNA Research Institute, 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan

W. C. Earnshaw  
Wellcome Trust Centre for Cell Biology, University of Edinburgh, Edinburgh EH9 3JR, Scotland, UK

## Introduction

Gene therapy requires the provision of a wild-type copy of a gene to a gene-deficient cell in order to restore the proper expression and function of the damaged gene. Key requirements for successful gene therapy include: (1) the use of a full-length gene (or groups of genes) with all its regulatory elements, (2) a safe and effective vector to deliver the therapeutic gene to specific patient cells, and (3) stable long-term gene expression in a physiologically regulated fashion.

At present, full-length genes are available in characterized clones from human BAC and YAC libraries, or alternatively a desired gene can be selectively isolated from human genomic DNA using the transformation-associated recombination (TAR) cloning procedure, which exploits the high level of homologous recombination in yeast cells [1–3]. Human artificial chromosomes or HAC-based vectors are gene-delivery vectors with promising advantages

over currently used viral vectors [4–12]. All HACs by definition contain a functional centromere and can therefore replicate and segregate like normal chromosomes in human cells, without the need for integration into the host genome. Also, HAC vectors have essentially unlimited cloning capacity and are therefore able to carry genes or genomic fragments up to several Mb in extent. Finally, because they do not encode foreign proteins, HAC vectors minimize adverse host immunogenic responses and the risk of cellular transformation. Several studies have demonstrated the use of HAC-based vectors for delivery and expression of genomic copies of genes in patient-derived gene-deficient cell lines and in mice ([10, 13–16] and references therein). The competence of HACs for transgenesis is favored by the presence of a unique recognition site for site-specific recombinases such as Cre, FLP, and phiC31 [17]. The first HAC vectors containing site-specific recombination sites for gene loading have been constructed using 21HAC engineered from the human chromosome 21 [10, 18].

One of the most versatile current HAC-based vectors is an alphoid<sup>tet<sup>O</sup></sup>-HAC that was engineered from a synthetic alphoid DNA array in human fibrosarcoma HT1080 cells [19]. This HAC comprises a circular megabase-size alphoid DNA array in which a single copy gene acceptor site was inserted [19–21]. The gene-loading site is within a centromere domain containing CENP-A nucleosomes interspersed with nucleosomes rich in histone H3 dimethylated at lysine 4 [22] and thus permissive for transcription [23]. Uniquely, this HAC can be easily eliminated from cell populations by inactivation of its conditional kinetochore [20, 24–27, 29]. HAC elimination may therefore be used to control for phenotypic changes attributed to expression of HAC-encoded genes [22]. The HAC remains structurally stable during months of propagation in host cells and does not integrate into the host chromosomes [21, 22]. Although the alphoid<sup>tet<sup>O</sup></sup>-HAC thus fulfills criteria for an appropriate gene-delivery vector, a detailed analysis of long-term gene expression from this HAC has been lacking.

For stable physiological expression of a gene of interest in centromere, the expressed gene should not disturb the centromere itself (because its disruption may induce inactivation of the HAC kinetochore followed by HAC loss) [19, 24]. Also, the inserted gene should evade silencing potentially induced as heterochromatin domains flanking centromere in all centromeric regions spread across the gene acceptor site [27, 28] (Fig. 1a). A logical strategy to overcome problems caused by shifting chromatin dynamics in the HAC centromere is to flank the transgene with insulator sequences as is routinely done for genes expressed from viral vectors [29–31].

In this study, to clarify a role of insulators in protecting gene expression from the HACs, we compared an effect of each of several candidate insulator sequences—cHS4,

gamma-satellite DNA, and a cluster of *tRNA* genes—on the expression of a *EGFP* transgene inserted into the unique gene loading site of the alphoid<sup>tet<sup>O</sup></sup>-HAC vector. Our results indicate that proximity to centromere does not protect genes loaded into the HAC vector from silencing if they are not flanked by insulator sequences. Our study also demonstrates that a well-known cHS4 insulator as well as two newly discovered insulators, tDNA and gamma-satellite DNA, provide strong protective effects on transgene expression in the HAC. In two analyzed host cell lines, the highest barrier activity was observed for a tDNA insulator consisting of two functional copies of *tRNA* genes. HAC-based vectors carrying a therapeutic gene protected by any of these insulator sequences should be useful in the development of optimal strategies for stable transgenesis in different types of cells.

## Materials and methods

### Cell culture

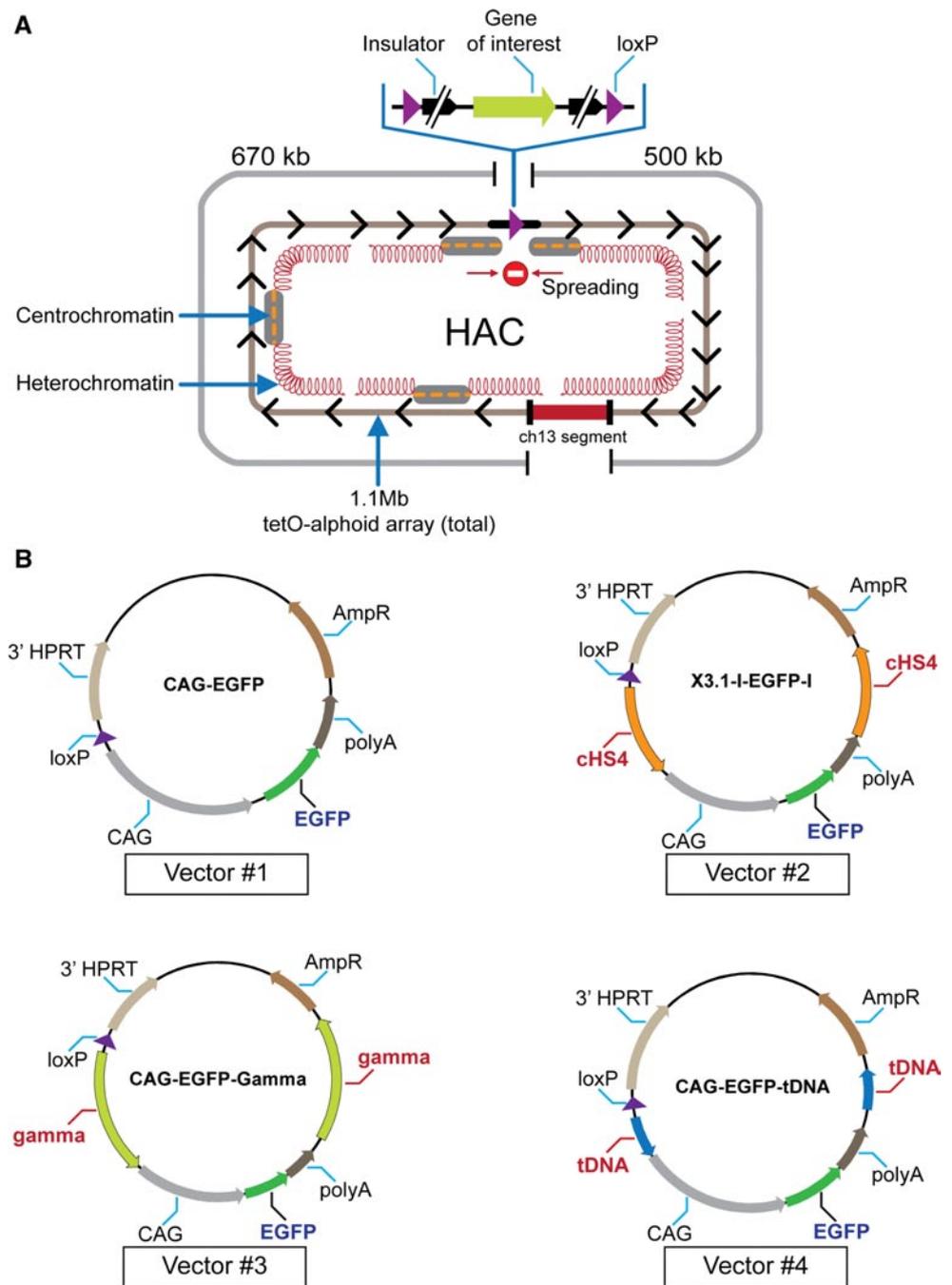
HeLa cells were cultured in Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Carlsbad, CA, USA) supplemented with 10 % fetal bovine serum (FBS; Atlanta Biologicals, Lawrenceville, GA, USA), 1 % penicillin–streptomycin (P/S; Invitrogen, USA) with 2 µg/ml of Blasticidin S (BS; Invitrogen) at 37 °C in 5 % CO<sub>2</sub>. Hypoxanthine phosphoribosyltransferase (HPRT)-deficient Chinese hamster ovary (CHO) cells (JCRB0218) were maintained in Ham's F-12 nutrient mixture (Invitrogen) plus 10 % FBS, 1 % P/S and 8 µg/ml of BS. After introduction of the *EGFP*-containing vectors, CHO cells retaining the alphoid<sup>tet<sup>O</sup></sup>-HAC-EGFP were cultured with 1 × HAT medium (Invitrogen).

### Construction of the targeting vectors

The targeting vectors used in this study were constructed using standard ligation methods. Vectors #1 (CAG-EGFP), #3 (CAG-EGFP-Gamma), and #4 CAG-EGFP-tDNA were built in this study. Vector #2 (X3.1-I-EGFP-I) containing the green fluorescent protein (*EGFP*) transgene flanked on either sides by the *cHS4* insulator was built previously [20].

The vectors containing *EGFP* flanked by either tDNA or gamma-satellite DNA (gamma-satellite DNA from the human chromosome 8) [35] were constructed as follows. Two fragments of 500 bp each containing two copies of the *tRNA* gene were amplified from plasmid Rev-pCR4-TOPO [34] using a pair of primers SacI-tDNA/SpeISphI-tDNA (fragment A) and BstEII/BamHI-tDNA/NotI-tDNA (fragment B) (Table S1) and then cut by *SacI/SpeI* and *BstEII/NotI*, correspondingly. Vector X3.1-I-EGFP-I was digested with *SacI/SpeI* and ligated with fragment A and

**Fig. 1** **a** Diagram of the  $\text{aloid}^{\text{tetO}}$ -HAC with a unique gene loading  $\text{loxP}$  site [20]. The HAC contains <6,000 copies of the 42-bp tetracycline operator ( $\text{tetO}$ ) sequence incorporated into every second  $\text{aloid}$  DNA monomer of the 1.1 mega-base size  $\text{aloid}$  DNA array [19, 21]. Because  $\text{tetO}$  sequences are bound with very high affinity and specificity by the tet repressor ( $\text{tetR}$ ), the  $\text{tetO}$  sequences in the HAC can be targeted efficiently in vivo with  $\text{tetR}$  chromatin modifier fusion proteins. This can result in HAC elimination from cell populations by inactivation of its conditional kinetochore [19]. A gene-loading site is inserted into a transcriptionally permissive centrochromatin domain that is flanked by heterochromatin domains essential for centromere function. A repositioning of a heterochromatin domain(s) may potentially lead to silencing of genes loaded into the HAC. **b** Diagrams of the vectors used in this study. All vectors contain an  $\text{EGFP}$  color marker under a control of  $\text{CAG}$  promoter and a 3' HPRT- $\text{loxP}$  module. In three vectors (#2, #3, and #4),  $\text{EGFP}$  is flanked either by  $\text{cHS4}$ , gamma-satellite DNA or tDNA sequences, correspondingly. Vector #1 has no insulator sequence. The vectors were inserted into the  $\text{loxP}$  site of the  $\text{aloid}^{\text{tetO}}$ -HAC propagated in HPRT-deficient hamster CHO cells



then digested with *BstEII/NotI* and ligated with fragment B, producing the final vector CAG-EGFP-tDNA (#4). The gamma-satellite DNA was isolated as two fragments of 1.9 kb each by digestion of the vector gamma8-TOPO2.1 [35] by either *BamHI/NotI* or *SacI/SphI*. The *BamHI/NotI* fragment was ligated with CAG-EGFP-tDNA vector digested by *BamHI/NotI*. The resulting vector was then digested by *SacI/SphI* and ligated to the gamma-satellite DNA *SacI/SphI* fragment, producing the final vector CAG-EGFP-Gamma (#3).

The control vector CAG-EGFP (#1) containing the  $\text{EGFP}$  marker not flanked by insulators was constructed as follows. The basic vector X3.1-I-EGFP-I was digested by *EcoRI/NotI* and the fragment containing the 3'HPRT- $\text{loxP}$  module was isolated from the gel (fragment A). Next, a *EcoRI/NotI* linker (Table S1) was ligated with fragment A, producing the intermediate vector 261ENot. A fragment of 1,768 bp containing the  $\text{CAG}$  promoter was excised from the X3.1-I-EGFP-I vector by *KpnI/PacI* digestion and ligated with the 261ENot vector digested by *KpnI/PacI*,

producing the intermediate vector ENotCAG. A fragment of 1,273 bp containing the *EGFP* marker and the bovine globuline pA signal was PCR-amplified from the X3.1-I-EGFP-I vector using specific primers (Table S1) and then ligated with the ENotCAG vector digested by *PacI/NotI*, producing the final vector CAG-EGFP. This vector contains a 3'HPRT-loxP module and the *EGFP* marker under the *CAG* promoter. In vectors #2, #3 and #4, both 5' and 3' insulators that flank the *CAG-EGFP* gene were verified by sequencing using primers that anneal outside of the insulators.

#### FISH analysis of CHO cells with the BAC probe

FISH analysis was performed as previously described [20, 21]. HAC-containing CHO cells were cultured in F12 medium with 10  $\mu$ g/ml of colcemid (Invitrogen) overnight at 37 °C. Metaphase cells were trypsinized and centrifugated for 4 min at  $172 \times g$ , treated in 10 ml of 50 mM KCl hypotonic solution for 20 min at 37 °C and washed three times in methanol:acetic acid (3:1) solution with a 4-min centrifugation at  $172 \times g$  between each wash. Cells were diluted to the appropriate density with fixative solution, spread onto pre-cleaned slides (Thermo Fisher Scientific, Waltham, MA, USA) above steam (boiling water), and allowed to age 2 days at room temperature. For BAC probing, CHO metaphase slides were then washed in 70 % formamide in  $2 \times$  SSC for 2 min at 72 °C. Samples were dehydrated through a 70, 90, and 100 % ethanol series for 4 min each and left to air-dry. Orange 552 dUTP (5-TAMRA-dUTP) (Abbott Molecular, Des Plaines, IL, USA) labeled probes were denatured in hybridization solution at 78 °C for 10 min and left at 37 °C for 30 min. The hybridization mix probe was applied to the sample and incubated at 37 °C overnight. Slides were washed with  $0.4 \times$  SSC, 0.3 % Tween 20 for 2 min at 72 °C, briefly rinsed with  $2 \times$  SSC, 0.1 % Tween 20 (10 s) and air-dried in darkness. The samples were counterstained with VECTASHIELD mounting medium containing DAPI (Vector Laboratories, Burlingame, CA, USA). Slides were analyzed by fluorescence microscopy. Images were captured using a DeltaVision microscopy imaging system in the CRC, LRBGE Fluorescence Imaging Facility (NIH) and analyzed using ImageJ software (NIH). The probe used for FISH was BAC32-2-mer(tetO) DNA containing 40 kb of alphoid-tetO array cloned into a BAC vector as described previously [19]. BAC DNA was labeled using a nick-translation kit with Orange 552 dUTP (5-TAMRA-dUTP) (Abbott Molecular).

#### FISH analysis of HeLa cells with PNA probes

Preparation of HeLa metaphase slides was identical to that of CHO metaphase slides with the exception that HeLa

metaphase cells were obtained by overnight incubation in DMEM medium with 0.05  $\mu$ g/ml colcemid (Invitrogen). For PNA probing, slides were rehydrated with  $1 \times$  PBS for 15 min, and fixed in 4 % formaldehyde- $1 \times$  PBS for 2 min, followed by three 5-min washes in  $1 \times$  PBS, then dehydrated with a 70, 90, and 100 % ethanol series for 5 min each, and left to air-dry. PNA (peptide nucleic acid) labeled probes used were telomere (CCCTAA)<sub>3</sub>-Cy3 (PerSeptive Biosystems, Framingham, MA, USA) and tetO-alphoid array (FITC-OO-ACCACTCCCTATCAG) (Panagene, South Korea). Ten nanomoles of each probe were mixed with hybridization buffer and applied to each slide. Slides were then incubated at 80 °C for 3 min and hybridized for 2 h at room temperature in darkness. Slides were washed twice in 70 % formamide, 10 mM Tris-HCl; pH 7.2, 0.1 % BSA, followed by three washes with  $1 \times$  TBS, 0.08 % Tween-20. Slides were then dehydrated through an ethanol dehydration series, air-dried, and counterstained with VECTASHIELD mounting medium containing DAPI (Vector Labs). Images were captured using a DeltaVision microscopy imaging system in the CRC, LRBGE Fluorescence Imaging Facility (NIH) and analyzed using ImageJ software (NIH).

#### Microcell-mediated chromosome transfer

Microcell-mediated chromosome transfer (MMCT) was performed as described previously [20, 22]. The alphoid<sup>tetO</sup>-HACs containing the 3' HPRT/loxP/EGFP/Insulator vectors were transferred from CHO (Chinese hamster ovary) into HeLa (human cervical carcinoma) cells via MMCT. Briefly, microcells were prepared by centrifugation of  $1 \times 10^9$  CHO cells attached to flasks (Nalge Nunc, USA) coated with poly-L-lysine (Sigma-Aldrich, St. Louis, MO, USA) and fused to  $1 \times 10^6$  HeLa cells using a Neo EX HVJ Envelope Transfection Kit (Cosmo Bio, Carlsbad, CA, USA). HeLa hybrids were selected in the presence of 2  $\mu$ g/ml of blasticidin (BS) and collected for expansion.

#### Genomic DNA preparation and PCR analysis

Genomic DNA for PCR analysis was prepared using DNAzol (Invitrogen). Reconstitution of the *HPRT* gene after Cre/lox-mediated recombination was confirmed by primers Lox137-R and Rev #6. Cross contamination by CHO chromosomes was quantified using hamster-specific primers, Furin-F and Furin-R. Human-specific primers dog-F1 and dog-R1 were used to verify that the fused cells were human. All these primers are listed in Table S1.

#### FACS analysis

Analysis of EGFP fluorescence on CHO and HeLa cells was performed using an LSR Fortessa flow cytometer (BD

Biosciences, Franklin Lakes, NJ, USA), with the data processed by BD FACSDiva software (BD Biosciences). A minimum of  $1 \times 10^4$  cells was analyzed for each sample. The intensity of EGFP fluorescence was calibrated using a Rainbow Calibration Particle (6.0–6.4  $\mu\text{m}$ ) standard (Spherotech, Inc., Lake Forest, IL, USA).

#### ChIP assay and real-time PCR

ChIP with antibodies against trimethyl H3 Lys4 (Millipore, Bedford, MA, USA) was carried out using a chromatin immunoprecipitation (ChIP) assay kit (Millipore) according to the manufacturer's instructions. Each sample used  $1.2 \times 10^6$  cultured cells, which were cross-linked in 1 % formaldehyde (Electron Microscopy Sciences-EMS, Hatfield, PA, USA) for 10 min at 37 °C. Soluble chromatin was prepared by two 5-min sonications at 8 °C (Bioruptor sonicator; Cosmo Bio) in lysis buffer with added protease inhibitors, yielding an average DNA size of 500 bp. We used 2  $\mu\text{l}$  per sample of H3K4me3, with 2.5  $\mu\text{l}$  of mouse IgG per sample used as a control for unspecific binding. The recovery ratio of the immunoprecipitated DNA relative to input DNA was measured by real-time PCR using a 7900HT Fast Real-Time PCR system (Applied Biosystems, Carlsbad, CA, USA) and Syber Green PCR master mix (Applied Biosystems). The results within each experiment were normalized to unspecific binding of mouse IgG antibody, while comparisons between different insulators was made by normalizing to rDNA, which was assumed to be unaffected by the presence of insulators in the HAC. Primers for *EGFP*, *bsr*, *Hygro*, and *HSV-TK* transgene sequences as well as for control loci, 5S ribosomal DNA and pericentromeric satellite 2 repeats are listed in Supplementary Table S1. At least three independent ChIP experiments were performed to estimate the level of enrichment.

#### Statistical analysis

Statistical analysis was made using Prism (GraphPad Software Inc., La Jolla, CA, USA). The data were normalized using a logarithmic transformation and a two-way ANOVA was conducted to examine the effect of the insulator type and culturing time on the intensity of EGFP fluorescence of the cell. A post hoc Tukey test was performed if a significant effect was found and  $p < 0.05$  was considered significant.

#### TSA and SAHA treatment

Clones without insulators (N1, N5, and N7 for HeLa cells and n4 for CHO cells) that have been cultured for 12 weeks were grown on a six-well plate. Once 90 % confluence was reached, the cells were incubated in 2 ml of nonselective

medium containing either 100 ng/ml of Trichostatin A (TSA) (Wako) or 10 mM of suberoylanilide hydroxamic acid (SAHA) (Vorinostat) for 24 h. These cultures were then collected and analyzed by FACS.

#### Results

The insulators we have tested here include the most known and widely used one, cHS4, discovered by Felsenfeld and collaborators [32]. cHS4 is derived from DNase hypersensitive site 4 of the locus control region of the chicken  $\beta$ -globin gene cluster. cHS4 exhibits both enhancer-blocking and barrier activity, making it a good candidate for use in gene transfer applications. An important characteristic of the cHS4 insulator is that it functions in a vast number of vertebrate cell types and organisms tested, including rodent and human cells [29]. Other insulators have been less commonly used for experiments on gene expression in mammalian cells. Among them are two recently identified chromatin insulators, tDNA and gamma-satellite DNA, which are proposed to be essential elements in the establishment and maintenance of chromatin domain architecture in mammalian cells [33–35]. It was demonstrated that a cluster of *tRNA* genes can protect a randomly integrated transgene from silencing in human and mouse cells [33, 34]. Similarly, short tandem arrays of gamma-satellite repeats present in pericentromeric regions of all human chromosomes exhibit a strong anti-silencing activity [35]. Previously we exploited a reporter gene assay for barrier element function in mouse MEL cells [36] to demonstrate that both *tRNA* genes and gamma-satellite DNA can protect local chromatin structure by preventing heterochromatin spreading from flanking regions [34, 35].

#### Construction of vectors with the *EGFP* transgene flanked by cHS4, tDNA, or gamma-satellite DNA insulators

Four vectors were used for comparative analysis of insulator activity (Fig. 1b). All vectors contain the green fluorescent protein (*EGFP*) transgene and a 3' HPRT-loxP module designed for site-specific targeting into the single loxP gene loading site of the alphoid<sup>tet<sup>O</sup></sup>-HAC using Cre/lox mediated recombinase. Vector #2 (X3.1-I-EGFP-I), in which the transgene is flanked by the full-length 1.2 kb cHS4 element on both sides, was constructed previously [29]. Two other newly constructed vectors, CAG-EGFP-Gamma (#3) and CAG-EGFP-tDNA (#4), contain the *EGFP* transgene flanked by either a 0.5-kb segment containing two *tRNA* genes (*Gly* and *Glu*) [34] or a 1.9-kb segment containing nine copies of 220-bp gamma-satellite repeat units from chromosome 8 [35]. The vector CAG-EGFP (#1) in which the *EGFP* transgene is not flanked by insulators was also

constructed and used as a control. (For a detailed description of vector construction see "Materials and methods").

#### Insertion of the vectors into a single loxP site of the alphoid<sup>tetO</sup>-HAC in hamster CHO cells

The alphoid<sup>tetO</sup>-HAC containing a single loxP site [20] propagated in hamster HPRT-deficient CHO cells was used in these experiments. Each vector and a Cre-recombinase expression plasmid were co-transfected into CHO cells carrying the alphoid<sup>tetO</sup>-HAC (Fig. 2a). Integration of the vector into the loxP loading site within the HAC reconstitutes the *HPRT* gene, allowing cell selection on HAT medium. The recombinant clones were isolated by growth in HAT medium for 2–3 weeks. PCR analysis (Table S1) showed that the *HPRT* gene was indeed reconstituted in all drug-resistant clones (Fig. 2b). The efficiency of vector targeting into the HAC was  $\sim 1\text{--}2 \times 10^{-4}$ . All transfectants expressed the *EGFP* transgene. One EGFP-HAC clone (n4), one cHS4/EGFP-HAC clone (c11), two tDNA/EGFP-HAC clones (t1, t2), and two gamma/GFP-HAC clones (g1, g2) were identified by FISH analysis of CHO metaphase spreads and confirmed to carry autonomous HACs (Fig. 2d). These clones then were used for further analyses.

#### MMCT transfer of four different alphoid<sup>tetO</sup>-HACs to human HeLa cells

To examine the insulator activity of cHS4, gamma-satellite DNA and tDNA in human cells, HACs carrying the *EGFP* transgene flanked by different insulators (c1, g2, t2) and without insulator sequences (n4), were transferred from hamster CHO to human HeLa cells via microcell-mediated chromosome transfer (MMCT) (Fig. 2a). HeLa cells were chosen for these experiments because the level of heterochromatin-specific proteins in these cells is higher compared to other cell lines, which leads to fast silencing of transgenes [37]. As the alphoid<sup>tetO</sup>-HAC contains the blasticidin selectable marker (*bsr*) [19], selection of clones was carried out with BS. Between 3 and 7 drug-resistant clones that expressed the *EGFP* transgene were isolated for each HAC. PCR analyses with a set of hamster-specific primers (Table S1) showed that there was no cross contamination by the hamster chromosomes after MMCT (Fig. 2c). FISH analysis showed that the HACs were maintained autonomously without any detectable integration into the host genome in three out of five for EGFP-HAC (N1, N5, N7), in two out of seven clones for cHS4/EGFP-HAC vector (C4, C6), in three out of three clones for gamma/GFP-HAC vector (G1, G3, G10) and in three out of five randomly selected clones for tDNA/EGFP-HAC vector (T3, T5, T10) (Fig. 2e). These clones were used for further analyses.

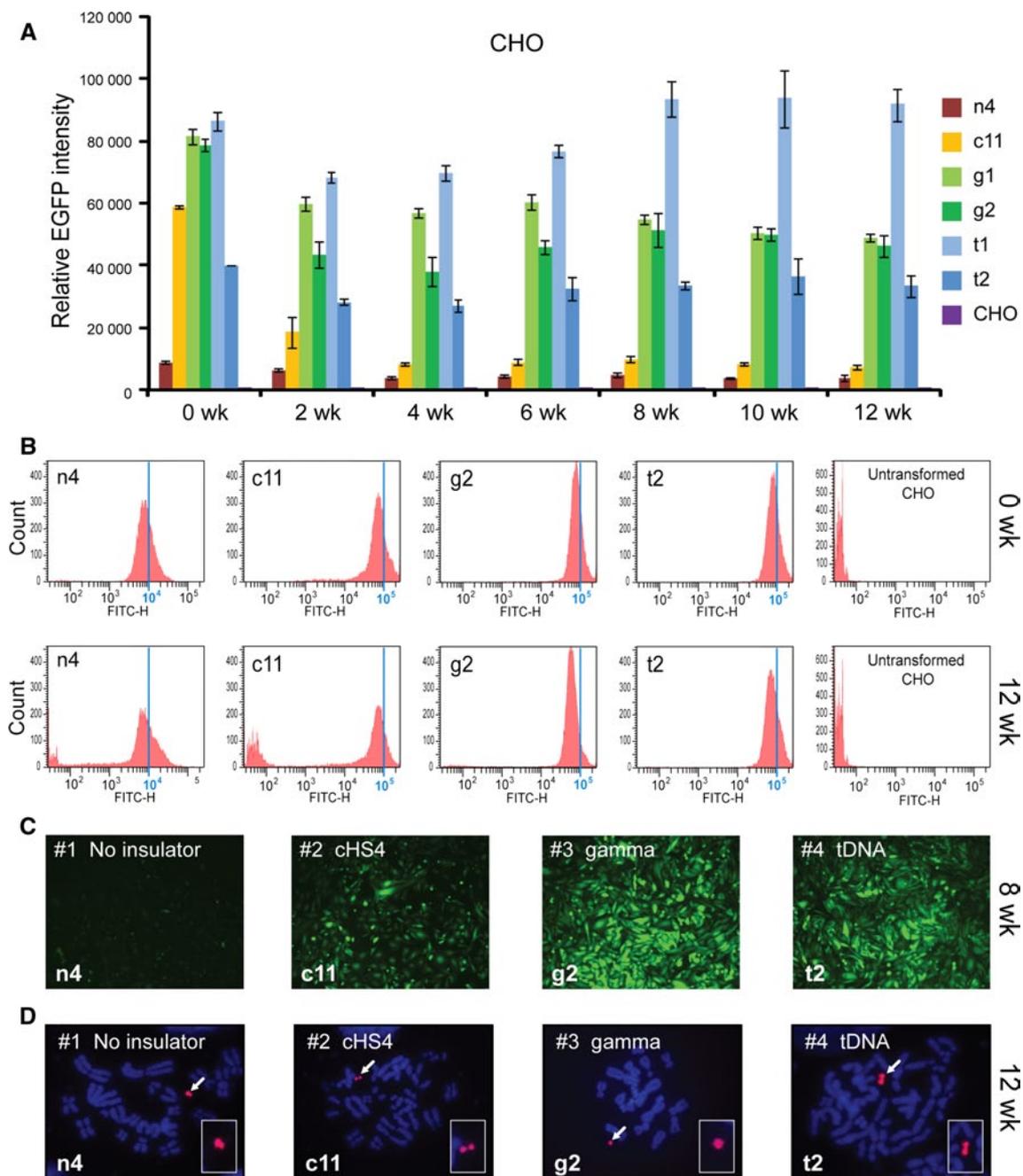
**Fig. 2** Schematic diagram of the generation of isogenic hamster and human cell lines containing the alphoid<sup>tetO</sup>-HACs carrying four different constructs. **a** Steps of MMCT to transfer a HAC into hamster CHO and then to human HeLa cells. The alphoid<sup>tetO</sup>-HAC was engineered in human HT1080 cells [19]. The loxP gene loading site was inserted into the HAC by homologous recombination in chicken DT40 cells [20]. This modified HAC was transferred to hamster CHO cells. In those cells, four vectors containing the *EGFP* transgene with or without flanking insulator elements and a 3' HPRT-loxP module were inserted into the loxP site of the HAC by Cre-loxP mediated recombination, producing isogenic hamster cell lines. From CHO cells, autonomously propagated HACs carrying four different vectors were transferred to HeLa cells, producing isogenic human cell lines. **b** Loading of the vectors into the loxP site of the HAC propagated in HPRT-deficient hamster CHO cells is accompanied by reconstitution of the *HPRT* gene allowing cell selection on HAT medium. Lanes 1, 2, 3, and 4 correspond to PCR products obtained with the genomic DNA isolated from HAC-containing clones (n4, c11, g2 and t2) of CHO cells that were chosen for MMCT transfer to human HeLa cells. **c** PCR control for cross contamination of human HeLa cells by hamster chromosomes. Lanes 5–10 correspond to randomly chosen clones of HeLa cells containing the HAC (G10, T5, C6); clones were checked by PCR using furin-F/furin-R hamster specific primers (lanes 5–7) or by PCR using dog-F1/dog-R1 human specific primers (lanes 8–10). Lanes 11, 12 correspond to the HAC-containing clone of HeLa cells (N5) checked by PCR using either hamster (lane 11) or human (lane 12) specific primers. FISH analysis of **d** the hamster clones (n4, c11, g2, t2) carrying the autonomously propagated HACs and **e** human clones (N5, C6, G10, T5) carrying the autonomously propagated HACs. FISH analysis in human cells was performed using PNA labeled probes for telomeric (red) and tetO-alphoid sequences (green). FISH analysis in hamster cells was performed using the BAC probe specific for the HAC-vector sequence (red) (see details in "Materials and methods"). Chromosomal DNA was counterstained with DAPI (blue). The HACs are indicated by arrows

#### Comparative analysis of EGFP expression flanked by different insulator sequences

The *EGFP* transgene expression from the HACs could be monitored by fluorescence microscopy. The level of *EGFP* transgene expression from each HAC construct was quantitated using fluorescence activated cell sorting (FACS) analysis by measuring the geometric mean green fluorescence of the cell population. Two cell lines, hamster CHO and human HeLa, carrying different HACs, were used for analysis. Each measurement was the average of values of three independent experiments for each clone. EGFP expression was measured every 2 weeks for 3–5 months.

For CHO cells, one clone with the EGFP-HAC construct (n4), one clone with the cHS4/EGFP-HAC construct (c11), two clones with the gamma/GFP-HAC construct (g1, g2), and two clones with the tDNA/EGFP-HAC construct (t1, t2), were analyzed by FACS. A two-way ANOVA was conducted in order to examine the effect of the insulator type and culturing time on the EGFP intensity. At the start of the experiment (0 week), all insulator constructs exhibited significantly higher EGFP expression than the control (n4) (Fig. 3a) [two-way ANOVA,  $F(3, 98) = 369.7$ ; 0 week





**Fig. 3** Long-term *EGFP* transgene expression in hamster CHO cells. **a** Relative mean *EGFP* fluorescence determined by FACS of cells carrying different HAC constructs grown under BS selection (clones n4, c11, g1, g2, t1, t2). Error bars represent standard deviation. *EGFP*-positive cells were sorted and analyzed by FACS over the course of 3 months, with measurements taken every 2 weeks. The measurement for each clone is the average of three experiments. **b** Flow cytometry histograms of levels of *EGFP* fluorescence in the population at the beginning of the experiment (0 weeks) and after 12 weeks of culture

under BS selection. The x-axis represents the intensity of the fluorescence, the y-axis the number of cells. **c** Fluorescence images of CHO cells with the *alploid<sup>tetO</sup>*-HAC carrying the different *EGFP* constructs after 8 weeks (n4, c11, g2, t2) of culture. Cells were cultured in BS<sup>+</sup> medium to select for retention of the HAC. **d** FISH analysis of representative CHO clones (n4, c11, g2, t2) carrying the autonomously propagated HACs after 12 weeks of culture under selection in BS<sup>+</sup> medium. Chromosomal DNA was counterstained with DAPI (blue). The HACs are indicated by arrows

by 2.5-fold over the course of the 12-week experiment (Fig. 3a). This decrease in *EGFP* intensity became significant after 10 weeks of culturing (n4, 0 vs. 10 week,

$p = 0.0131$ ). The gradual emergence of non-fluorescence cells during the experiment was the cause of this decrease in the average *EGFP* intensity in the cell population of the

clone c11. Figure 3b shows representative flow cytometry histograms of the level of EGFP fluorescence in cultures at the beginning of the experiment (0 weeks) and after 12 weeks of culture under selection for BS. Fluorescence images of CHO cells with the different alphoid<sup>tet<sup>O</sup></sup>-HAC vectors (n4, c11, g2, t2) are shown in Fig. 3c after 8 weeks of culture. FISH analysis demonstrated that the HACs remained in an autonomous form in all clones analyzed after 12 weeks of culture under the HAC selection (Fig. 3d).

For HeLa cells, three clones for EGFP-HAC (N1, N5, N7), two clones for cHS4/EGFP-HAC construct (C4, C6), three clones for gamma/GFP-HAC construct (G1, G3, G10), and three clones for tDNA/EGFP-HAC construct (T3, T5, T10), were used for analysis. As seen from Fig. 4a, at the start of the experiment (0 week) all constructs, including those with the cHS4 insulator sequence, exhibited significantly higher EGFP expression compared to the control [two-way ANOVA;  $F(3, 313) = 495.5$ ; 0 week cHS4,  $p < 0.0001$ ; 0 week gamma,  $p < 0.0001$ ; 0 week tDNA,  $p < 0.0001$ ]. The expression level of all insulator constructs did not change significantly after 12 weeks of culture under BS selection [two-way ANOVA;  $F(6, 313) = 1.808$ ,  $p = 0.0969$ ]. Neither did we observe any significant change in the EGFP fluorescence in clones C4, C6, G1, G10, T5, and T10 that were cultured longer, up to 22 weeks under BS selection [two-way ANOVA,  $F(11, 179) = 2.261$ ,  $p = 0.1303$ ]. Figure 4b shows representative flow cytometry histograms of the level of EGFP fluorescence in cultures at the beginning of the experiment (0 weeks) and after 20 weeks of culture. Unlike in CHO, no significant population of EGFP(-) cells emerged in HeLa clones with any of the three insulator constructs. Fluorescent images of one representative clone for each construct, i.e., tDNA/EGFP-HAC (T5), gamma/GFP-HAC (G10), cHS4/EGFP-HAC (C6), and EGFP-HAC (N5), are shown in Fig. 4c. FISH analysis showed that the HACs are maintained autonomously in the host genome in all clones analyzed after culturing for 20 weeks with the HAC selection (in BS<sup>+</sup> medium) (Fig. 4d).

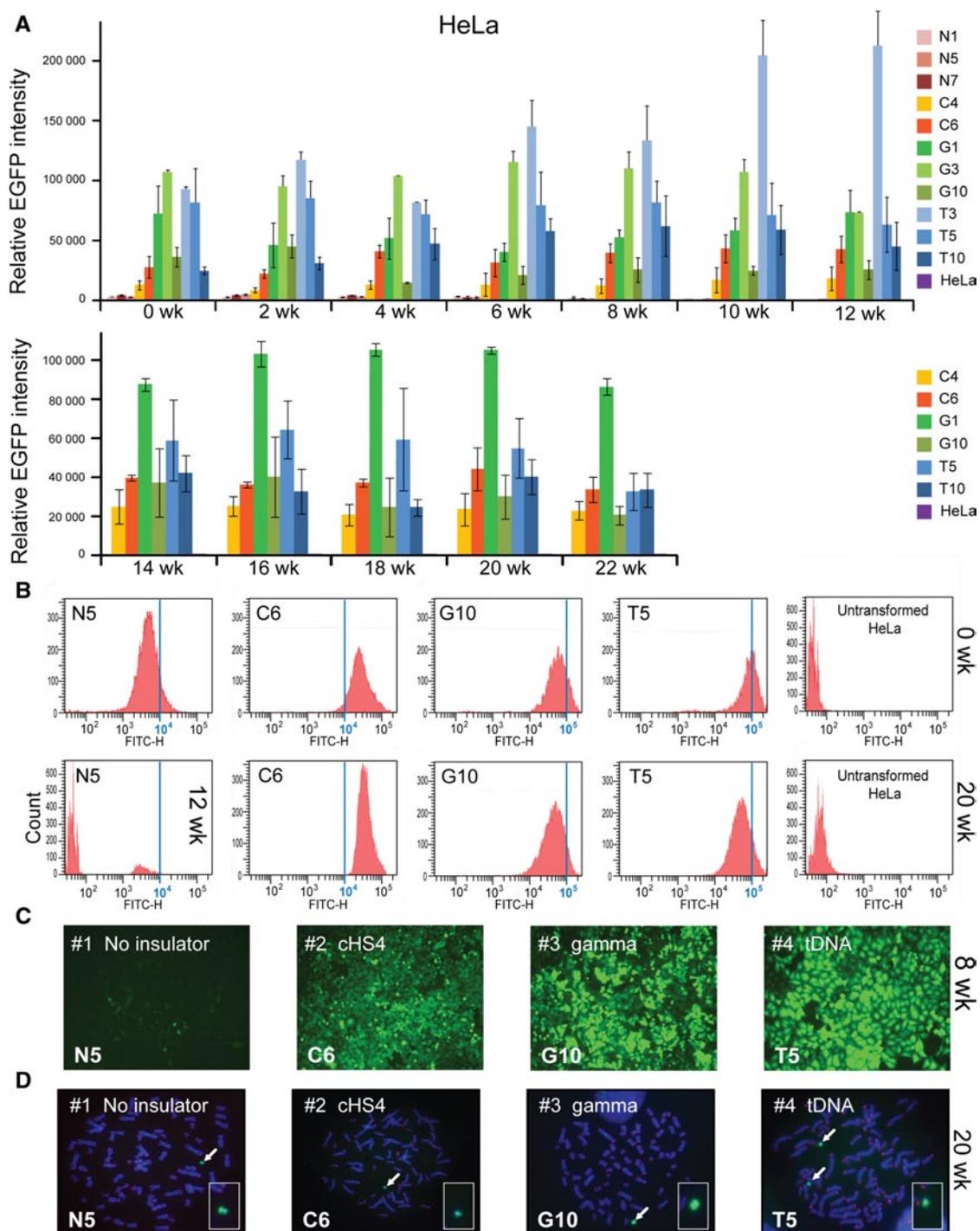
The average (geometric mean) EGFP intensity over the duration of the 12-week experiment in CHO cells was approximately 3-, 11-, and 11-fold higher for the HAC constructs carrying either cHS4, gamma-satellite DNA or tDNA sequences, respectively, compared to the control construct without *EGFP* flanking sequences (Fig. 5a). This difference in expression between the control and insulators was significant [two-way ANOVA,  $F(3, 98) = 369.7$ ; cHS4,  $p < 0.0001$ ; gamma,  $p < 0.0001$ ; tDNA,  $p < 0.0001$ ]. There was no significant difference in the EGFP intensity between HAC constructs carrying gamma-satellite DNA or tDNA ( $p = 0.9659$ ). However, both gamma-satellite DNA ( $p < 0.0001$ ) and tDNA ( $p < 0.0001$ )-containing HAC

constructs were significantly more intense in their EGFP fluorescence than cHS4/EGFP-HAC, despite no initial differences at the start of the experiment. The difference in EGFP intensity between insulators in CHO emerged after 2 weeks of culture (cHS4 vs. gamma,  $p = 0.0002$ ; cHS4 vs. tDNA,  $p = 0.017$ ).

In HeLa cells, the average (geometric mean) EGFP intensity over 12 weeks of culture was approximately 11-, 24-, and 35-fold higher for the HAC constructs carrying either cHS4, gamma-satellite DNA or tDNA sequences respectively compared to the control constructs without EGFP flanking sequences (Fig. 5b) [two-way ANOVA,  $F(3, 313) = 495.5$ ; cHS4,  $p < 0.0001$ ; Gamma,  $p < 0.0001$ ; tDNA,  $p < 0.0001$ ]. In HeLa cells, the HAC constructs carrying tDNA had significantly higher EGFP fluorescence than gamma-satellite DNA constructs ( $p = 0.0001$ ). Both gamma-satellite DNA ( $p < 0.0001$ ) and tDNA ( $p < 0.0001$ )-containing HAC constructs were also significantly brighter than cHS4/EGFP-HAC. In the control cell lines without insulators (N1, N5, N7), a growing population of EGFP(-) cells began to emerge at 6 weeks of culturing. This caused a significant decrease in the average EGFP intensity of the cell population by 12 weeks [two-way ANOVA;  $F(6, 313) = 1.808$ ,  $p < 0.0001$ ] (Fig. 5c).

To summarize, the level of the EGFP expression during a 3-month period in hamster (CHO) and a 5-month period in human cells (HeLa) was highest when the *EGFP* transgene was flanked by either gamma-satellite DNA or tDNA sequences. In both cell types, the presence of insulators significantly increased the intensity of the EGFP fluorescence.

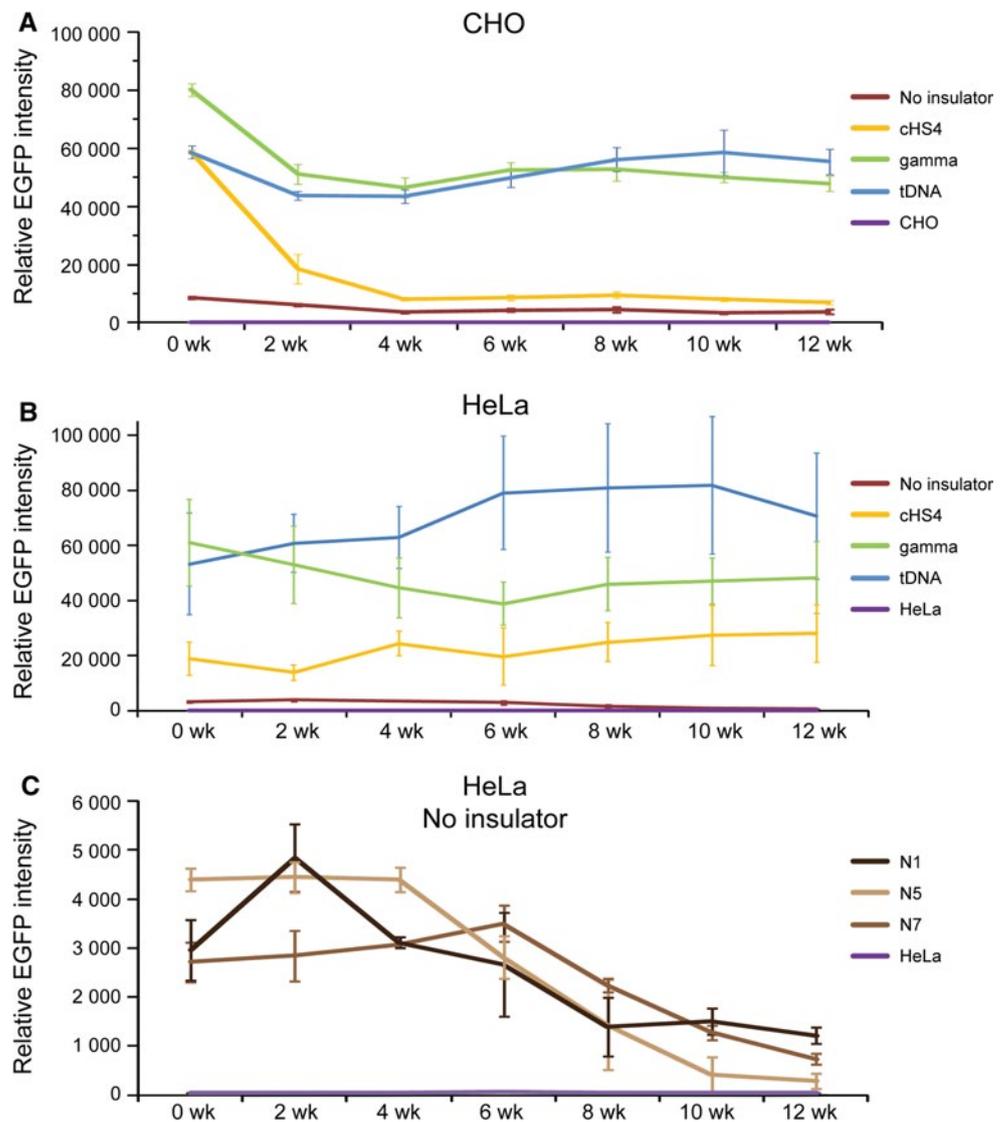
Chromatin immunoprecipitation (ChIP) analysis was employed to analyze the chromatin structure of the alphoid<sup>tet<sup>O</sup></sup>-HAC DNA carrying different cassettes in human HeLa cells. For this experiment, we examined three HAC clones (C6, G1, T5) after 19 weeks of culture and one HAC clone (N5) after 6 weeks of culture. More specifically, the HAC region containing the single copy *EGFP* transgene, *Hygro*, and *HSV-TK* genes as well as the *bsr* gene (which was repeated ~30–40 times within the mega-size alphoid DNA array) [19] were analyzed (Fig. 6a). ChIP analysis revealed that H3K4me3, a marker for active gene promoters, was highly associated only with the *EGFP* transgene sequence flanked by cHS4 or tDNA or gamma-satellite DNA sequences (Fig. 6b). A very low association was observed for the *bsr* gene sequences scattered throughout the megabase-size alphoid array of the HAC as well as for *Hygro* and *HSV-TK* genes not protected by insulator sequences. We assume that the lower levels on the *bsr* genes reflect the fact that most copies of this gene are likely to be inactive on the HAC. These results clearly indicate that cHS4, tDNA, as well as gamma-satellite DNA sequences work as barrier elements protecting a transgene



**Fig. 4** Long-term *EGFP* transgene expression in human HeLa cells. **a** Relative mean *EGFP* fluorescence determined by FACS of cells carrying different HAC constructs grown under BS selection similar to that described in Fig. 3 for hamster CHO cells. Each data point represents the mean  $\pm$  SD from three independent experiments. Three clones for the HAC construct without insulator (N1, N5, N7), two clones for the HAC construct with chicken insulator cHS4 (C4, C6), three clones for the HAC construct with gamma-satellite DNA (G1, G3, G10) and three clones for the HAC construct with tDNA (T3, T5, T10) were cultured during 12 weeks. Two clones for each HAC construct (C4, C6, G1, G10, T5, T10) were grown for 22 weeks. Cells

were cultured with HAC selection (in BS<sup>+</sup> medium). **b** Flow cytometry histograms of the levels of *EGFP* fluorescence in the population at the beginning of the experiment (0 weeks) and after 20 weeks of culture. The *x*-axis represents the intensity of the fluorescence, the *y*-axis the number of cells. **c** Fluorescence images of HeLa cells with the alphoid<sup>tetO</sup>-HAC carrying the different *EGFP* transgenes after 8 weeks of culture (N5, C6, G10, T5 clones). **d** FISH analysis of representative HeLa clones (N5, C6, G10, T5) carrying the autonomously propagated HACs after 20 weeks of culture under selection in BS<sup>+</sup> medium. Chromosomal DNA was counterstained with DAPI (blue). The HACs are indicated by arrows

**Fig. 5** Temporal analysis of the geometric mean EGFP intensity of clones of CHO (a) and HeLa (b) cells with the  $\text{alphoid}^{\text{tetO}}$ -HAC carrying different *EGFP* transgenes (brown no insulator; yellow with cHS4; green with gamma-satellite DNA; blue with tDNA; purple untransformed cells) during 12 weeks of culture with selection in  $\text{BS}^+$  medium. c EGFP expression undergoes a progressive decline over the course of 12 weeks if the gene is not flanked by insulator sequences in HeLa cells. The error bars represent the pooled standard deviation of all clones from each insulator type



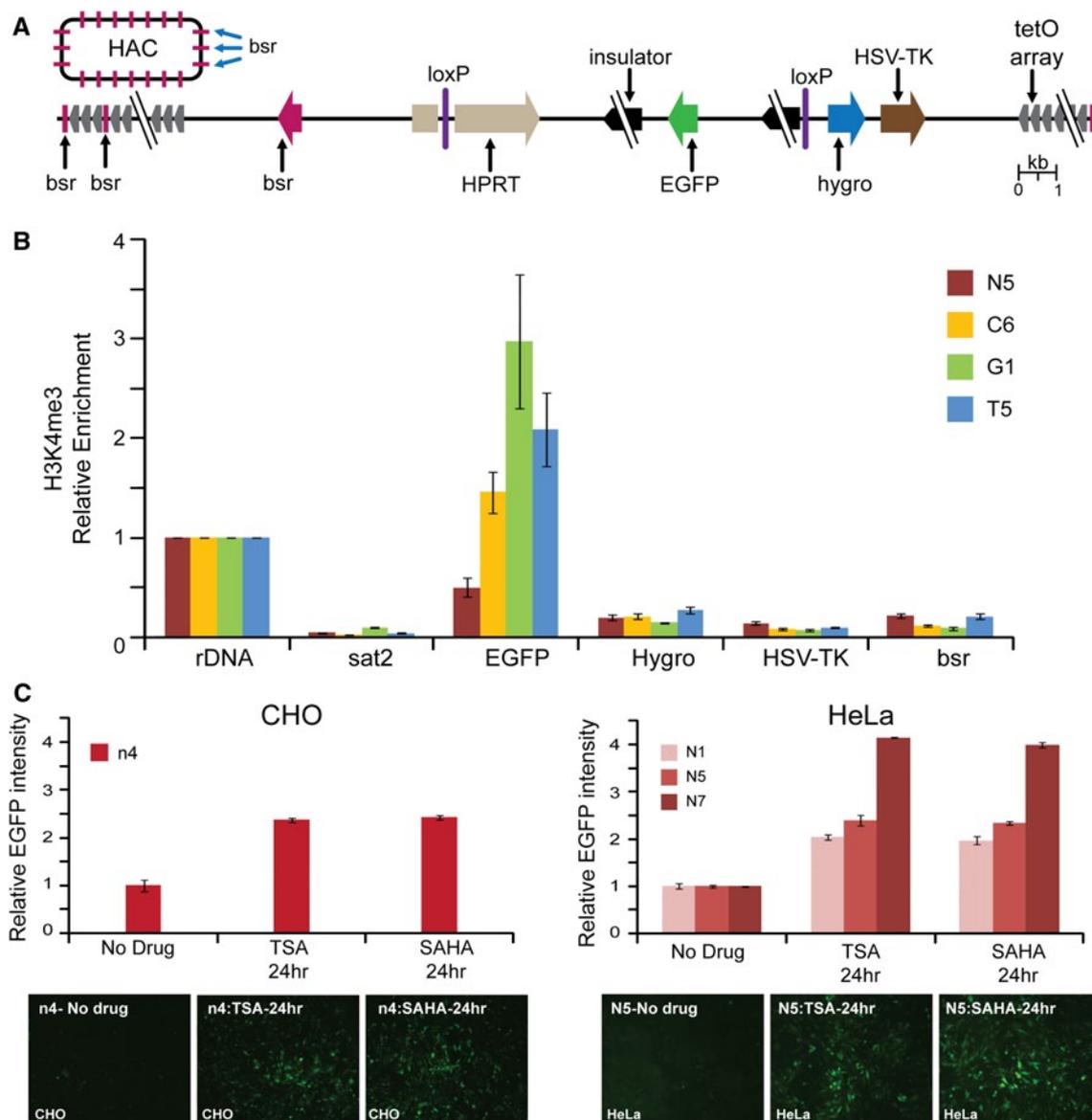
against heterochromatin domains on the  $\text{alphoid}^{\text{tetO}}$  DNA array of the HAC. A statistical difference in enrichment of H3K3me3 on EGFP between insulators was determined by one-way ANOVA. A Tukey post hoc test revealed that enrichment of H3K4me3 on EGFP in the C6 clone (cHS4) was statically significantly lower than in either G1 (gamma-satellite DNA) [ $F(3, 12) = 50.81, p < 0.0001$ ] or in T5 (tDNA) ( $p = 0.0274$ ) clones. There were no statistically significant differences between gamma-satellite DNA and tDNA ( $p = 0.557$ ). All three insulator constructs were significantly higher than the control (N7) (C6,  $p = 0.0002$ ; G1,  $p < 0.0001$ ; T5,  $p < 0.0001$ ). This pattern is similar to that of the FACS results.

*EGFP* gene silencing in control clones N1, N5, and N7 in HeLa cells and in the control clone n4 in hamster CHO cells is likely to be caused by heterochromatinization of the *EGFP* region initiated from nearby  $\text{alphoid}$  DNA

sequences of the HAC. Indirect support for this conclusion comes from experiments where these clones containing the *EGFP* transgene without insulator protection were treated with the histone deacetylase (HDAC) inhibitors, TSA (trichostatin), and SAHA (suberoylanilide hydroxamic acid) [38]. EGFP expression was reactivated when these inhibitors were added to the culture (Fig. 6c, d), suggesting that an increase in acetylation level of histone H3 in the HAC favors transcription of the transgene.

## Discussion

The expression of transferred exogenous genes is subject to the influence of surrounding chromatin. The phenomenon of “chromosomal position effects” frequently leads to rapid and irreversible transgene silencing. Several studies have



**Fig. 6** ChIP analysis of H3K4me3 chromatin in the transgene cassettes on the alphoid<sup>tetO</sup> HAC propagated in HeLa cells. **a** A map of the transgenes within the HAC; HSV-TK, Hygro, HPRT genes plus an EGFP color marker flanked on both sides by different insulators. The bsr gene is repeated 30–40 times within the 1.1 Mb tetO-alphoid array of the alphoid<sup>tetO</sup>-HAC [19, 21]. Arrows indicate the direction of transcription of the transgenes. **b** ChIP analysis was performed for clones C6, G1, T5, and N5. Enrichment is shown relative to the 5S rRNA control locus. Satellite 2 sequence, Sat2, from endogenous pericentromeric repeats was included as a negative control. **c** The n4 clone of CHO and N1, N5, and N7 clones of HeLa cells were treated either with TSA or SAHA. EGFP intensity is shown relative to the non-treated cultures. Fluorescence images of cells with and without

drug treatment are shown. In the CHO control cell line (n4), a significant increase in GFP intensity was observed upon treatment with either TSA or SAHA [one-way ANOVA,  $F(2, 15) = 589.0$ ; TSA,  $p < 0.0001$ ; SAHA,  $p < 0.0001$ ]. There was no significant difference in GFP intensity between TSA- and SAHA-treated cells ( $p = 0.4523$ ). A similar pattern was observed in HeLa control cell lines N1, N5, and N7. GFP expression was significantly increased in HeLa cells upon drug treatment. [one-way ANOVA,  $F(8, 27) = 1287$ ; N1-TSA,  $p < 0.0001$ ; N1-SAHA,  $p < 0.0001$ ; N5-TSA,  $p < 0.0001$ ; N5-SAHA,  $p < 0.0001$ ; N7-TSA,  $p < 0.0001$ ; N7-SAHA,  $p < 0.0001$ ]. No significant difference was detected between TSA- and SAHA-treated cells (N1,  $p = 0.4615$ ; N5,  $p = 0.9597$ ; N7,  $p = 0.0502$ )

shown that flanking a transgene with chromatin insulators can reduce both the rate and extent of transgene silencing and position effect variegation [29–31].

The cHS4 insulator derived from the chicken  $\beta$ -globin locus has been intensively used to protect transgene expression [32]. Barrier activity of the native cHS4 element is

associated with a peak of histone hyperacetylation and other hallmarks of an “open” histone modification profile. These are thought to prevent the invasion of transcriptionally repressive heterochromatin from an adjoining region into the chicken  $\beta$ -globin locus [39]. Subsequent studies demonstrated that this activity is associated in part with the binding of a protein complex, including USF1 and other histone modifying enzymes, at specific points within and adjacent to the cHS4 core region [40].

A similar mechanism, but involving other protein complexes, has been proposed for two other insulators identified in human genomes, tDNA and gamma-satellite DNA. The barrier activity of tDNA requires binding of Pol III factors and their interaction with chromatin remodeling and histone modifying enzymes in a spatio-temporal organization of these loci in a nuclear zone [33, 34]. Interestingly, *tRNA* genes have long been known to flank the central core of centromeric DNA in *S. pombe* [41].

In contrast to *tRNA* genes scattered in the human genome, gamma-satellite DNA is localized exclusively in the pericentromeric regions of all human chromosomes. Gamma-satellite DNA arrays at their natural location may act to prevent heterochromatin spreading beyond the pericentromeric regions. Anti-silencing activity of gamma-satellite insulator depends on the binding of Ikaros [35], which promotes chromatin remodeling and activation of target genes in pericentromeric chromatin [42].

Based on accumulating literature data, a level of gene “protection” by an insulator depends both on the nature of the insulator sequence and on the chromatin organization of the regions neighboring the insulator [30, 31]. Therefore, the effect of insulators on expression of genes loaded into de novo formed HAC vectors cannot be absolutely predicted. In the case of the alphoid<sup>tet<sup>O</sup></sup>-HAC, our previous chromatin analyses indicated that our synthetic gene-loading site is inserted into a transcriptionally permissive centromeric domain [22, 28]. In the HACs as well as in centromeres of normal chromosomes, this domain is flanked by heterochromatin domains essential for centromere function [43]. Based on such organization, two scenarios could be considered. First, centromeric chromatin may represent a privileged region for Pol II-transcribed genes, as was previously shown for rice centromeres [44]. In this case, insulators may not be required at all for stable gene expression in this region. Alternatively, genes located in centromeric chromatin may undergo silencing—caused either by specific features of centromeric chromatin that support only a very low level of transcription on the underlying centromeric repeat sequences [23, 26] or by repositioning of heterochromatin domain(s) after replication of centromeric DNA.

In this study, using a set of isogenic cell lines carrying the alphoid<sup>tet<sup>O</sup></sup>-HAC vector, we investigated the effect

of three chromatin insulators, cHS4, and gamma-satellite DNA and tDNA, on the expression of the *EGFP* transgene. Our results indicate that epigenetic silencing of the unprotected transgene occurs when it is in close proximity to centromeric chromatin. Formally, this is in agreement with both (i) the hypothesis of spreading of heterochromatin domain(s) within a functional centromere, and (ii) with a repressive effect of centromeric chromatin domain. We prefer the first alternative based on the kinetics of extinction of expression of the insulator-less transgene. A low level of EGFP expression observed just after gene loading progressively declined during subsequent cell culturing (Fig. 5c), as might be expected if there is a gradual heterochromatinization of the transgene sequence (but not if centromeric chromatin represses transgene expression from the outset).

With all three insulators, transgene expression remained at a relatively high level, with no evidence of significant silencing in human cells. The highest barrier activity was detected for a tDNA insulator consisting of two functional copies of *tRNA* genes and gamma-satellite DNA repeats. It is worth noting that a similar effect was observed when these chromatin insulators were analyzed at an ectopic chromosomal site in mouse cells [34, 35]. Thus, chromatin insulators positioned in centromeric chromatin function similarly to their action previously reported at other chromosomal locations.

The strong protective effect of insulators observed in this study indicates that chromatin insulators should be included in future experiments on long-term gene expression in HACs. These chromatin insulators could be inserted into the HAC along with a gene of interest. For example, insulators can be included in the retrofitting vector pJBRV1 [22] that has been specifically constructed for loading full-length genes isolated by transformation-associated recombination (TAR) cloning [1–3] into the alphoid<sup>tet<sup>O</sup></sup>-HAC. Because the activity of an insulator may vary in different cells, it would be practically useful to flank a gene of interest with different insulators and test which is more suitable for a specific host cell environment. Notably, the HAC-based vector described here may also be used to discover new barrier elements that prevent gene silencing in centromeric chromatin.

**Acknowledgments** This work was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research, USA (V.L. and N.K.), the Wellcome Trust of which W.C.E. is a Principal Research Fellow (grant number 073915), the Grand-in-Aid for Scientific Research from Ministry of Education, Culture, Sports, Science and Technology of Japan (H.M.), the Kazusa DNA Research Institute Foundation (H.M.), by Goskontrakt 02.512.11.2085 (A.V.T.) and by the programme of the Russian Academy of Sciences « Molecular and Cellular Biology » (A.V.T.). A.V.T., E.T., and M.L. were also supported by the Program of the Presidium of the Russian Academy of Sciences “Molecular and Cellular Biology”. The authors would like to thank the CRC, LRBGE

Fluorescence Imaging Facility (NIH), and personally Drs. Karpova and Dr. McNally for instructions, consultations and help with the usage of a DeltaVision microscopy imaging system.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Larionov V, Kouprina N, Graves J et al (1996) Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc Natl Acad Sci USA* 93:491–496
- Kouprina N, Larionov V (2006) TAR cloning: insights into gene function, long-range haplotypes and genome structure and evolution. *Nat Rev Genet* 7:805–812
- Kouprina N, Larionov V (2008) Selective isolation of genomic loci from complex genomes by transformation-associated recombination cloning in the yeast *Saccharomyces cerevisiae*. *Nat Protocols* 3:371–377
- Harrington JJ, Van Bokkelen G, Mays RW et al (1997) Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. *Nat Genet* 15:345–355
- Saffery R, Choo KH (2002) Strategies for engineering human chromosomes with therapeutic potential. *J Gene Med* 4:5–13
- Basu J, Willard HF (2006) Human artificial chromosomes: potential applications and clinical considerations. *Pediatr Clin North Am* 53: 843–53, viii
- Monaco ZL, Moralli D (2006) Progress in artificial chromosome technology. *Biochem Soc Trans* 34(Pt 2):324–327
- Ren X, Tahimic CG, Katoh M et al (2006) Human artificial chromosome vectors meet stem cells: new prospects for gene delivery. *Stem Cell Rev* 2:43–50
- Oshimura M, Katoh M (2008) Transfer of human artificial chromosome vectors into stem cells. *Reprod Biomed* 16:57–69
- Kazuki Y, Oshimura M (2011) Human artificial chromosomes for gene delivery and the development of animal models. *Mol Ther* 19:1591–1601
- Ikeno M, Suzuki N (2011) Construction and use of a bottom-up HAC vector for transgene expression. *Methods Mol Biol* 738:101–110
- Kazuki Y, Hiratsuka M, Takiguchi M et al (2010) Complete genetic correction of iPSCs cells from Duchenne muscular dystrophy. *Mol Ther* 18:386–393
- Kouprina N, Earnshaw WC, Hiroshi Masumoto H, Larionov V (2013) A new generation of human artificial chromosomes for functional genomics and gene therapy. *Cell Mol Life Sci* 70:1135–1148
- Tedesco FS, Hoshiya H, D'Antona G et al (2011) Stem cell-mediated transfer of a human artificial chromosome ameliorates muscular dystrophy. *Sci Transl Med* 3:96ra78
- Tedesco FS, Gerli MF, Perani L et al (2012) Transplantation of genetically corrected human iPSC-derived progenitors in mice with limb-girdle muscular dystrophy. *Sci Transl Med* 4:140ra89
- Kazuki Y, Kobayashi K, Aueviriyavit S, Oshima T, Kuroiwa Y, Tsukazaki Y, Senda N, Kawakami H, Ohtsuki S, Abe S, Takiguchi M, Hoshiya H, Kajitani N, Takehara S, Kubo K, Terasaki T, Chiba K, Tomizuka K, Oshimura M (2012) Trans-chromosomal mice containing a human CYP3A cluster for prediction of xenobiotic metabolism in humans. *Hum Mol Genet* 22:578–592
- Branda CS, Dymecki SM (2004) Talking about a revolution: the impact of site-specific recombinases on genetic analyses in mice. *Dev Cell* 6:7–28
- Yamaguchi S, Kazuki Y, Nakayama Y, Nanba E, Oshimura M, Ohbayashi T (2011) A method for producing transgenic cells using a multi-integrase system on a human artificial chromosome vector. *PLoS ONE* 6:e17267
- Nakano M, Cardinale S, Noskov VN et al (2008) Inactivation of a human kinetochore by specific targeting of chromatin modifiers. *Dev Cell* 14:507–522
- Iida Y, Kim JH, Kazuki Y et al (2010) Human artificial chromosome with a conditional centromere for gene delivery and gene expression. *DNA Res* 17:293–301
- Kouprina N, Samoshkin A, Erliandri I et al (2012) Organization of synthetic alphoid DNA array in human artificial chromosome (HAC) with a conditional centromere. *ASC Synthetic Biology* 1:590–601
- Kim JH, Kononenko A, Erliandri I et al (2011) Human artificial chromosome (HAC) vector with a conditional centromere for correction of genetic deficiencies in human cells. *Proc Natl Acad Sci USA* 108:20048–20053
- Chan FL, Marshall OJ, Saffery R, Kim BW, Earle E, Choo KH, Wong LH (2012) Active transcription and essential role of RNA polymerase II at the centromere during mitosis. *Proc Natl Acad Sci USA* 109:1979–1984
- Cardinale S, Bergmann JH, Kelly D et al (2009) Hierarchical inactivation of a synthetic human kinetochore by a chromatin modifier. *Mol Biol Cell* 20:4194–4204
- Bergmann JH, Rodríguez MG, Martins NM et al (2011) Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. *EMBO J* 30:328–340
- Bergmann JH, Jakubsche JN, Martins NM et al (2012) Epigenetic engineering: histone H3K9 acetylation is compatible with kinetochore structure and function. *J Cell Sci* 125(Pt 2): 411–421
- Bergman JH, Martins NMC, Larionov V et al (2012) HACKing the centromere chromatin code: in sights from human artificial chromosomes. *Chrom Research* 20:505–519
- Schueler MG, Sullivan BA (2006) Structural and functional dynamics of human centromeric chromatin. *Annu Rev Genomics Hum Genet* 7:301–313
- Felsenfeld G, Burgess-Beusse B, Farrell C, Gaszner M, Ghirlando R, Huang S, Jin C, Litt M, Magdinier F, Mutskov V, Nakatani Y, Tagami H, West A, Yusufzai T (2004) Chromatin boundaries and chromatin domains. *Cold Spring Harb Symp Quant Biol* 69:245–250
- Wallace JA, Felsenfeld G (2007) We gather together: insulators and genome organization. *Curr Opin Genet Dev* 17:400–407
- Barkess G, West AG (2012) Chromatin insulator elements: establishing barriers to set heterochromatin boundaries. *Epigenomics* 4:67–80
- Pikaart MJ, Recillas-Targa F, Felsenfeld G (1998) Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev* 12:2852–2862
- Raab JR, Chiu J, Zhu J, Katzman S, Kurukuti S, Wade PA, Haussler D, Kamakaka RT (2011) Human tRNA genes function as chromatin insulators. *EMBO J* 31:330–350
- Ebersole T, Kim JH, Samoshkin A, Kouprina N, Pavlicek A, White RJ, Larionov V (2011) tRNA genes protect a reporter gene from epigenetic silencing in mouse cells. *Cell Cycle* 10:2779–2791
- Kim JH, Ebersole T, Kouprina N, Noskov VN, Ohzeki J, Masumoto H, Mravinac B, Sullivan BA, Pavlicek A, Dovat S, Pack SD, Kwon YW, Flanagan PT, Loukinov D, Lobanenko V, Larionov V (2009) Human gamma-satellite DNA maintains open chromatin structure and protects a transgene from epigenetic silencing. *Genome Res* 19:533–544

36. Feng YQ, Lorincz MC, Fiering S, Grealley JM, Bouhassira EE (2001) Position effects are influenced by the orientation of a transgene with respect to flanking chromatin. *Mol Cell Biol* 21:298–309
37. Ohzeki JI, Bergmann JH, Kouprina N et al (2012) Breaking the HAC Barrier: histone H3K9 acetyl/methyl balance regulates CENP-A assembly. *EMBO J* 31:2391–2402
38. Chavan AV, Somani RR (2010) HDAC inhibitors—new generation of target specific treatment. *Mini Rev Med Chem* 10:1263–1276
39. Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G (2001) Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 293:2453–2455
40. West AG, Huang S, Gaszner M, Litt MD, Felsenfeld G (2004) Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell* 16:453–463
41. Takahashi K, Murakami S, Chikashige Y, Niwa O, Yanagida M (1991) A large number of tRNA genes are symmetrically located in fission yeast centromeres. *J Mol Biol* 218:13–17
42. Li Z, Perez-Casellas LA, Savic A, Song C, Dovat S (2011) Ikaros isoforms: the saga continues. *World J Biol Chem* 2:140–145
43. Pidoux AL, Allshire RC (2005) The role of heterochromatin in centromere function. *Philos Trans R Soc Lond B Biol Sci* 360:569–579
44. Nagaki K, Cheng Z, Ouyang S et al (2004) Sequencing of a rice centromere uncovers active genes. *Nat Genet* 36:138–145