The Sea Urchin *sns5* Insulator Protects Retroviral Vectors From Chromosomal Position Effects by Maintaining Active Chromatin Structure

Danilo D'Apolito[1,2], Elena Baiamonte[1], Mariella Bagliesi[1], Rosalba Di Marzo[1], Roberta Calzolari[1], Leda Ferro[3], Vito Franco[2], Giovanni Spinelli[3], Aurelio Maggio[1] and Santina Acuto[1]

[1]Unità di Ricerca "P. Cutino," U.O.C. Ematologia II, A.O. "V. Cervello," Palermo, Italy; [2]Dipartimento di Patologia Umana, Università di Palermo, Palermo, Italy; [3]Dipartimento di Biologia Cellulare e dello Sviluppo "Alberto Monroy," Università di Palermo, Palermo, Italy

Silencing and position-effect (PE) variegation (PEV), which is due to integration of viral vectors in heterochromatin regions, are considered significant obstacles to obtaining a consistent level of transgene expression in gene therapy. The inclusion of chromatin insulators into vectors has been proposed to counteract this position-dependent variegation of transgene expression. Here, we show that the sea urchin chromatin insulator, *sns5*, protects a recombinant γ-retroviral vector from the negative influence of chromatin in erythroid milieu. This element increases the probability of vector expression at different chromosomal integration sites, which reduces both silencing and PEV. By chromatin immunoprecipitation (ChIP) analysis, we demonstrated the specific binding of GATA1 and OCT1 transcription factors and the enrichment of hyperacetylated nucleosomes to *sns5* sequences. The results suggest that this new insulator is able to maintain a euchromatin state inside the provirus locus with mechanisms that are common to other characterized insulators. On the basis of its ability to function as barrier element in erythroid milieu and to bind the erythroid specific factor GATA1, the inclusion of *sns5* insulator in viral vectors may be of practical benefit in gene transfer applications and, in particular, for gene therapy of erythroid disorders.

Received 17 December 2008; accepted 13 March 2009; published online 7 April 2009. doi:10.1038/mt.2009.74

INTRODUCTION

Retroviral vectors are the most efficient means of stable gene delivery into hematopoietic stem cells. However, because of the integration into transcriptionally silent chromatin, a significant proportion of insertions are subjected to the repressive effects of the surrounding chromatin, which are termed chromosomal position effects (PEs), with silencing of provirus expression in a significant fraction of cells.[1] Chromatin PE can also be manifested as expression variegation (PEV), wherein, at a given point in time, genetically identical cells will show different phenotypes.[2,3] These adverse effects could be overcome to some extent by the incorporation of DNA elements into retroviral vectors, such as chromatin insulators, that function to establish and delimit domains of expression.[4–6] These elements, which were first described in *Drosophila* and are found in a wide range of organisms, can influence gene expression because of two experimentally defined properties, the positional enhancer blocker and barrier. The first function allows chromatin insulators to prevent promoter enhancer interactions only when placed between the two and, in doing so, can shield promoters from the influence of neighboring regulatory elements. In addition, by acting as barriers against the propagation of condensed chromatin, insulators may buffer a transgene from chromosomal PE.[7] Several studies have shown that the inclusion of the characterized cHS4 insulator in recombinant vectors reduces the rate and severity of vector silencing.[8–12] In addition, recent evidence has suggested that vectors that are insulated with this element may have a lower propensity to perturb nearby gene expression.[13] Therefore, the use of chromatin insulators is desirable in gene therapy approaches, as they have the ability to shield transgenes from the negative influence of chromatin, along with the potential to avoid insertional mutagenesis. Here, we investigated the ability of a new element, which is found in the early histone repeating unit of the sea urchin *Paracentrotus lividus*, to protect vector-encoded transgenes from chromosomal PE. We have previously identified an enhancer-blocker element of 265 bp, which has been termed *sns*, at the 3′ end of the *H2A* early histone gene that displays the capability to block enhancer-activated transcription in a polar and directional manner in both sea urchin and human cells.[14–16] This element functions as an enhancer blocker in erythroid milieu and binds erythroid and ubiquitous transcription factors.[17] Recently, studies in transgenic sea urchin embryos demonstrated that a 462 bp sequence, named *sns5,* that includes the *sns* sequence is necessary to regulate the transcription of the *H2A* early histone gene during sea urchin development, which suggests that the longer *sns5* sequence constitutes one of the borders of this transcription unit.[18] In light of these results, we have investigated the ability of both *sns* and *sns5* elements to prevent silencing and PEV on the expression of a γ-retrovirus vector. For this purpose, we have generated a large panel of mouse erythroleukemia (MEL) cells that bear integrated murine stem cell virus-based vectors that are flanked by insulators and are followed by the expression

Correspondence: *Santina Acuto, Unità di Ricerca "P. Cutino," Ematologia II, A.O. "V. Cervello," via Trabucco no. 180-90146, Palermo, Italy. E-mail: acutosanti@virgilio.it*

of a reporter-*GFP* gene. We found that *sns5*, but not *sns*, exhibits boundary properties when flanking the vector in a forward orientation. In order to better understand the mechanism that underlies the protection that is afforded by the *sns5* chromatin insulator in the setting of recombinant-viral vectors, we also used chromatin immunoprecipitation (ChIP) experiments to investigate the transcription factors and epigenetic modifications that are localized to the *sns5* insulator and the downstream long terminal repeat (LTR) promoter sequences. We show the colocalization of GATA1 and OCT1 transcription factors and the hyperacetylated nucleosomes to the *sns5* insulator, which suggests that this element is able to modify nucleosomal histones in order to maintain a euchromatic state inside the provirus locus.

RESULTS

Construction and characterization of reporter vectors

The maps of the DNA constructs that were used in these studies are shown in **Figure 1**. All constructs were derived from the stem cell retroviral vector, MGPN2.[19] We insulated the retroviral transgenes by inserting different DNA fragments into the vector 3′-LTR. PCR and sequence analysis on several transduced clones demonstrated that the inserted fragments were faithfully copied into the 5′-LTR after the retroviral replication (data not shown). The boundary properties of both *sns and sns5* were compared to that of the characterized 1.2 kb HS4 chromatin insulator from the chicken β-globin locus. As controls, the insulators were replaced by two spacers of different lengths (270 or 564 bp long) that were derived from λ-phage DNA. The inclusion of these various inserts had no significant effects on vector titer (**Table 1**) and stability, as shown by the appropriate size of provirus band in the Southern-blot analysis of selected (G418^R NIH-3T3) and unselected (MEL) individual clones (data not shown).

Frequency of vector expression in NIH-3T3 and MEL cells

In a preliminary analysis, we assessed the influence of the sea urchin elements *sns* and *sns5* in NIH-3T3 cells. G418 resistant clones (25–30 for each vector) from the higher dilution that was used for titration (to ensure low-copy integration) were expanded for 1 week, and the percentage of GFP$^+$ cells was measured by flow cytometry. We found a mean value of 56.8% ± 25.5 for the MGPN2 vector, 56.1% ± 23.8 for the *sns* (S) vector, and 68.9 ± 13 for the *sns5* (S) vector. These data indicated that the inclusion of the *sns5* element determined a higher mean percentage of GFP-expressing cells within the individual clones with a minor grade of variegation among clones as revealed from the minor value of SD. In light of this evidence, we switched to MEL cells, which have more pronounced vector silencing. In these experiments, the expression of the transgene was studied without selection to avoid any bias toward those subsets that are permissive for a threshold expression level, which are compatible with drug resistance.[20] MEL cells were transduced with the uninsulated and insulated vectors (**Figure 1**) using sets of ecotropic producer clones with similar vector titers, and they were subsequently cloned in a limiting dilution assay. Transductions were carried out at a modest multiplicity of infection in order to minimize the provirus copy number that was determined by Southern-blot analysis (data not shown). A large panel of MEL clones, which consisted of >600 unselected clones for each vector, derived from two independent transductions, was screened for vector integrants by multiplex PCR analysis using NEO specific primers for vector detection and mouse β-globin gene specific primers, as control. Flow cytometric analysis was performed for *GFP* expression in PCR$^+$ clones. For each vector, the percentage of clones that expressed the transgene was determined by considering positive clones to be those that have >5% fluorescent cells. The results have shown that the inclusion of the *sns5* element in a forward orientation increased the fraction of GFP-expressing cells to 85.8 ± 1.4%, as compared to the 45.2 ± 5.9% (P = 0.01) obtained with the uninsulated vector and the 56 ± 1.6% (P = 0.13) obtained with *sns5* in reverse orientation. The inclusion of the 1.2 kb cHS4 insulator in a forward orientation increased retroviral expression to 72.4 ± 6% (P = 0.04), which is

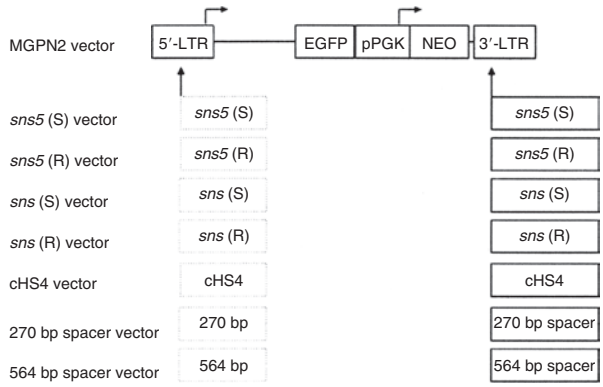

Figure 1 Vector constructs. The control vector MGPN2 based on the oncoretrovirus vector murine stem cell virus, contains two reporter cassettes: the *GFP* gene transcribed from the virus 5′-LTR promoter and the *NEO* gene transcribed from the PGK promoter. The insulated versions of this vector were constructed by inserting different DNA fragments into the 3′-LTR. In this location, the inserts are copied into the 5′-LTR to generate a flanking double copy as indicated by the dashed lines. The 462 bp DNA fragment *sns5*, from the 3′ end of the sea urchin *H2A* early histone gene, was inserted in forward and reverse orientation to the viral transcription to generate *sns5* (S) and *sns5* (R) vectors, respectively. Analogously, *sns* (S) and *sns* (R) vectors were generated by inserting the 265 bp *sns* element. The cHS4 vector contains the 1.2 kb HS4 insulator element from the chicken β-globin locus in forward orientation. Two fragments of 270 bp and 564 bp from λ-phage DNA were inserted to generate the 270 and 564 bp vectors, respectively. LTR, long terminal repeat.

Table 1 Reporter vector titers

Vector	Titer[a]
MGPN2	4–7 × 10^5
sns (S)	2–5 × 10^5
sns (R)	3–6 × 10^5
sns5 (S)	2–5 × 10^5
sns5 (R)	3–5 × 10^5
cHS4	1–3 × 10^5
Spacer 270 bp	3–7 × 10^5
Spacer 564 bp	2–4 × 10^5

[a]Range of vector titers of 10–15 GP+E86 producer clones expressed as G418^R colony-forming units/ml on naive NIH-3T3 cells.

a result that is similar to that reported by others using the same cHS4-insulated vector.[9] The *sns* (S) vector, *sns* (R) vector, 270 bp spacer vector, and 564 bp spacer vector did not show any statistically relevant effect on the frequency of expression (49.3 ± 11.5%, $P = 0.69$; 43.5 ± 6.5%, $P = 0.82$; 50.2 ± 3.2%, $P = 0.40$; and 50.03 ± 4.7%, $P = 0.46$, respectively). Detailed results are shown in **Table 2**. These results suggested that the presence of the *sns5* element, similar to the cHS4 element, increased the likelihood of the integrated retrovirus transgene to be expressed at different chromosomal positions. To further evaluate the ability of *sns5* to protect vector transgenes from the initial silencing effect, we measured the number of silenced proviruses 72 hours after transduction. We performed cell sorting experiments to separate cells that were expressing (GFP^+) and not expressing (GFP^-), the reporter gene, and determined the vector copy number per cell (VCN/cell), in both the positive and negative fractions. We found that a great number of uninsulated (MGPN2) provirus copies were silenced (average of 3.6 VCN/cell), whereas only 0.1 VCN/cell was detected in the GFP^- cells that were transduced with both the *sns5* (S) and cHS4 vectors. A mean of 0.6 VCN/cell and 3.5 VCN/cell was present in the *sns5* (R) and the *sns* (S) transduced cells, respectively (**Figure 2**). These results demonstrated the capability of the *sns5* element to improve the probability of vector expression in random integration sites and demonstrated that this element, like other insulators,[9] appears to function in the forward orientation only. Chromatin PE can also be manifested as PEV, wherein the progeny of a single cell can be affected by the spreading of the surrounding chromatin, which leads to a mosaic expression pattern.[21] To study the PEV, we expanded the GFP^+ pools of cells and measured the percentage of GFP^+ cells twice a week for 2 months. The analysis revealed a drop in the fraction of cells that were expressing the LTR→GFP cassette for all vectors during the first week of growth (**Figure 3**). The percentage of GFP^+ cells remained stable during the subsequent 50 days. The expression variegation that was detected within the first week of cell growth was quite different for the *sns5* (S) and cHS4 vectors when compared to the MGPN2, *sns* (S) and *sns5* (R) vectors. In particular, the reduction of GFP^+ cells was 10% for *sns5* (S), 16% for cHS4, 52% for MGPN2, 40% for *sns5* (R), and 35% for *sns* (S). Together, these data indicated that the *sns5* (S) and cHS4 insulators are both able to counteract the silencing of the integrated transgene immediately after transduction and during the first week of cell expansion. After that time, the extent of variegation did not change greatly in both insulated and uninsulated vectors, which suggested that most of the epigenetic modifications on the integrated transgene have already happened by that time and are subsequently inherited by the progeny. The PEV was also evaluated at the clonal level, 30 days after transduction, by measuring the percentage of GFP^+ cells in the progeny of cells that carried one or two provirus copies. A mean value of 13.1 ± 6.7% GFP^+ cells was found in clones that bore one copy of the uninsulated vector. Comparable values were detected in clones that were transduced with *sns* (S) (16 ± 9.2%) and *sns5* (R) (18.2 ± 12.2%) vectors. A higher percentage of GFP^+ cells was reported in the progeny of cells bearing a single copy of *sns5* (S) vector and cHS4 vector (79.9 ± 16.2% and 54.3 ± 25.7% respectively), with $P < 0.05$ (**Figure 4**). These results showed that the *sns5* insulator improved the likelihood of vector expression by counteracting both the initial silencing and the expression variegation. To confirm that the majority of *sns5* insulated provirus copies were transcriptionally active, the fluorescence intensity (MFC values) of each clone was plotted against the number of integrated transgenes. The results showed a direct correlation between the level of expression and

Table 2 Frequency of GFP expression in uninsulated and insulated vectors

Vector	% GFP^+/PCR^+[a]	VCN/cell[b]
MGPN2	49.4	1.4
	41.0	1.4
sns (S)	57.4	1.1
	41.2	0.9
sns (R)	39.0	0.9
	48.1	1.3
sns5 (S)	86.8	0.8
	89.8	1.1
sns5 (R)	57.1	1.4
	54.9	1.2
cHS4	76.7	0.8
	68.2	0.9
Spacer 270 bp	52.5	1.2
	47.9	1.1
Spacer 564 bp	46.7	1
	55.4	1.3

Abbreviations: GFP, green fluorescent protein; VCN, vector copy number.
[a]Percentage of the clones that express the reporter gene (*GFP* gene) in clones that are PCR^+ for provirus-specific sequences. [b]VCN per cell determined by Southern-blot analysis.

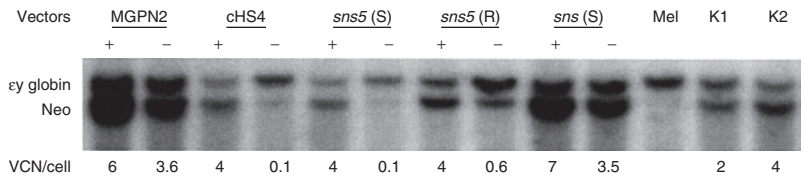

Figure 2 Quantification of gene transfer rates by Southern-blot analysis. Southern-blot analysis was performed on genomic DNA from sorted GFP^+ and GFP^- cells transduced with uninsulated and insulated vectors. DNA was digested with *Eco*RV restriction enzyme, which cuts once in each viral long terminal repeat, and the blot was hybridized contemporarily with specific the proviral probe (NEO sequence) and the specific mouse endogenous εγ globin probe. The DNA from untransduced cells was used as negative control. The band intensities were quantified with a phosphoimager, and provirus signals normalized with the εγ endogenous globin gene. DNAs from two clones bearing two copies (K1) and four copies (K2) of the provirus were loaded as a control for probe-binding efficiency. Vector copy number per cell (VCN/cell) is indicated at the bottom of each line.

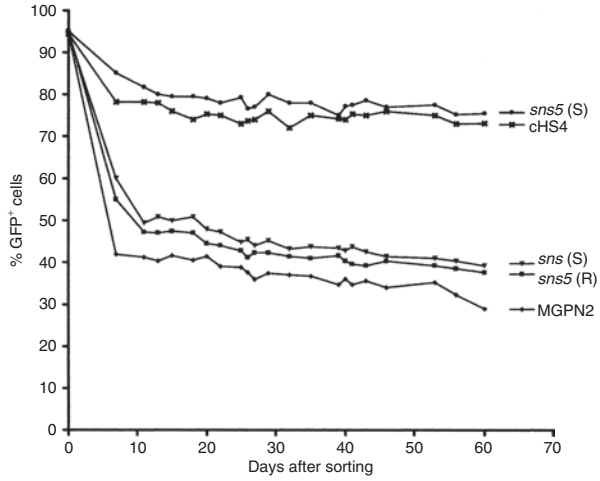

Figure 3 Extinction of *GFP* expression after cell sorting. The GFP$^+$ sorted fractions from mouse erythroleukemia cells transduced with the uninsulated (MGPN2) and insulated vectors (cHS4 vector, *sns5* (S) vector, *sns5* (R) vector, and *sns* (S) vector) were analyzed by flow cytometry, determining the percentage of GFP$^+$ cells twice a week for 60 days. The percentages of GFP$^+$ cells are shown on the *y*-axis and days of analysis on the *x*-axis.

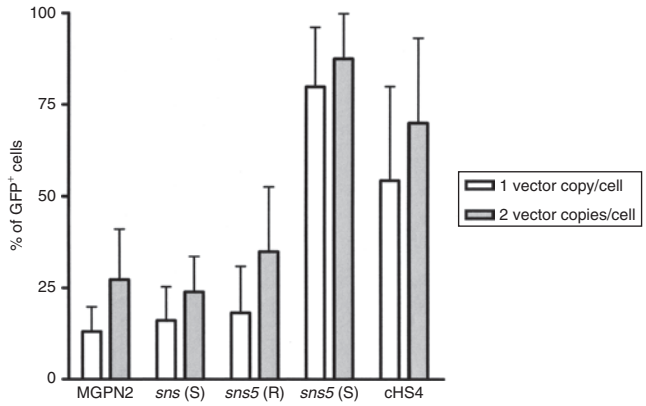

Figure 4 Percentages of GFP$^+$ cells in individual mouse erythroleukemia clones. The frequency of fluorescent cells was measured, 30 days after cell expansion, in individual clones bearing one or two copies of vector. In this analysis, 11–15 clones carrying one copy (white bars) and 18–24 clones carrying two copies (gray bars) of proviruses are included. The diagram represents the mean of GFP$^+$ cell percentages ± SD.

the VCN in cells that were transduced with the *sns5* (S) insulated vector (slope value: 0.52; *r* value: 0.89) and, to a lesser degree, with the cHS4 insulated vector (slope value: 0.22; *r* value: 0.54). No significant correlation was observed with any of the other vectors (MGPN2 vector, slope: 0.13; *r* value: 0.46; *sns5* (R) vector, slope: 0.06, *r* value: 0.41; *sns* (S) vector, slope: 0.051, *r* value: 0.37) (**Figure 5**). This analysis confirmed that all the vectors tested, except those that were insulated with *sns5* (S) and cHS4, are highly susceptible to transcriptional silencing that is indicated by a low transcription level, despite multiple integration events. These data also revealed that the mean level of *GFP* gene expression among clones containing one copy of vector did not vary significantly between the various vectors that were tested, which suggested that these chromatin insulators do not have stimulatory or inhibitory effects on their own.

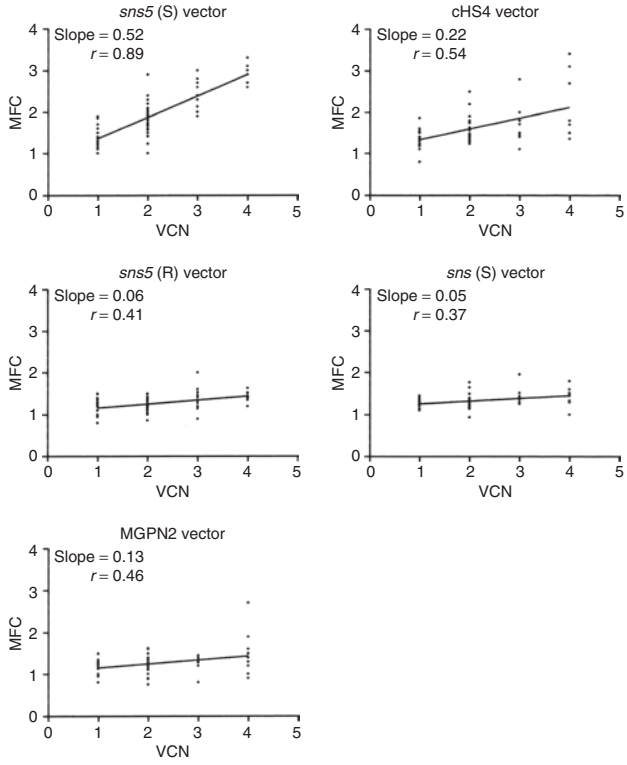

Figure 5 Effect of insulators on copy-number dependent expression. To check for a potential copy-number dependent effect after 30 days of cell expansion the mean fluorescent channel (MFC) values (*y*-axis) of the clones that contain uninsulated or insulated vectors were plotted with the number of integrated vectors (*x*-axis). The analysis were performed in clones bearing single or multiple integrated events with a minimum of six clones for each group. Linear regression analysis was used for statistical evaluation. VCN, vector copy number.

Transcription factors mapping and chromatin architecture of the *sns5* insulated vector

In order to elucidate the mechanism by which the sea urchin *sns5* insulator improves the expression of integrated transgene in MEL cells, we performed ChIP analysis on three GFP$^+$ and two GFP$^-$ clones that each carried a single copy of the *sns5* (S) vector. The integrated vectors were checked for the absence of mutations by sequencing overlapping PCR fragments that covered from the 5′-LTR to the *GFP* gene (data not shown). Initially, we assessed the effective association *in vivo* of the erythroid GATA1 and the ubiquitous OCT1 transcription factors that were previously demonstrated to bind *sns5 in vitro*.[17] Previous EMSA analysis had shown two GATA1 binding regions in the *sns5* element, which were in a sequence that is common to both the *sns5* and *sns* fragments[17] and a sequence that is located at the 3′ end of *sns5* fragment, which is not included in the shorter *sns* element (data not shown). Soluble chromatin samples were incubated with anti-GATA1 and anti-OCT1 antibodies. The specifically immunoprecipitated DNA was analyzed by quantitative real-time PCR on the same amount of immunoprecipitated and nonimmunoprecipitated (input) chromatin using Taqman probes for the *sns5* and 5′-LTR vector sequences. We found a mean of eightfold enrichment of the *sns5* sequence in the immunoprecipitated chromatin using both antibodies in GFP$^+$ clones (8.4 ± 2.3 for GATA1

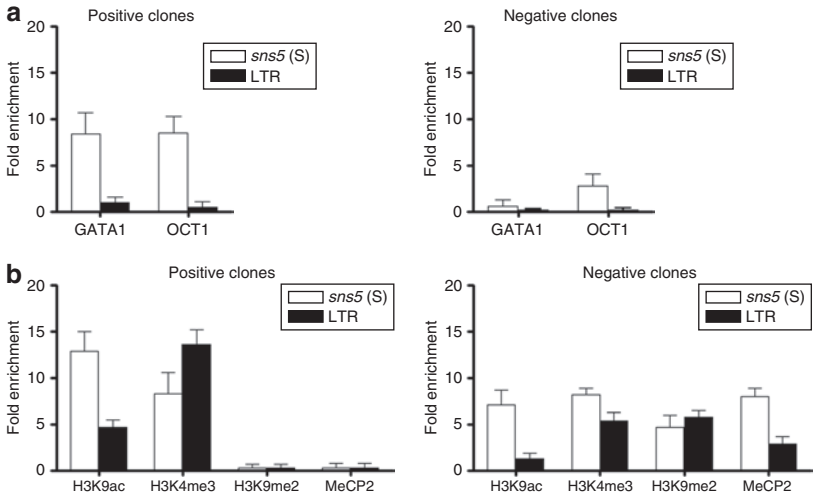

Figure 6 ***sns5*** **recruits** ***trans*****-acting factors and increases provirus histone acetylation and other markers of active chromatin** ***in vivo.*** ChIP analysis was performed in single-copy vector clones using antibody against GATA1 and OCT1 transcription factors (**a**) and antihistone modifications (**b**). Mouse erythroleukemia cells transduced with *sns5* (S) vector were subcloned and scored for the presence of provirus by PCR and for *GFP* gene expression. PCR$^+$ clones (GFP$^+$ and GFP$^-$ clones) were analyzed by Southern blotting for vector copy number and integrity. Clones were checked for the "double-copy" configuration and fidelity of retrotranscription process by sequencing of PCR fragment. ChIP analysis was performed on chromatin from three positive and two negative clones. Soluble chromatin, digested by micrococcal DNase to an average size of 200–600 bp, was immunoprecipitated with antibodies against GATA1 and OCT1 transcription factors and with antibodies against histone modifications marking open chromatin (H3K4me3), and (H3K9ac). Antibodies against the H3K9me2 and MeCP2 were used as markers of inactive chromatin. Immunoprecipitated DNA was quantitatively analyzed by real-time PCR using primers specific for *sns5* sequences (white bars) and for the LTR promoter region (black bars). The ratio between the amount of provirus-specific template before and after chromatin immunoprecipitation was determined by quantitative PCR in the same amount (5 ng) of chromatin. Histograms represent the average enrichment values ± SD (*y*-axis) of at least three independent ChIP experiments for each clone in which PCR was performed twice (in triplicate). In *x*-axis are represented the antibodies used. ChIP, chromatin immunoprecipitation; LTR, long terminal repeat.

and 8.4 ± 1.8 for OCT1) (**Figure 6a**). Little or no enrichment was detected for both transcription factors in transcriptionally inactive cells (0.5 ± 0.8 for GATA1 and 2.8 ± 1.3 for OCT1). As expected, little or no enrichment was observed in the LTR promoter region in both GFP$^+$ and GFP$^-$ clones, because no consensus motifs for GATA1 and OCT1 transcription factors are present in the retroviral LTR sequences. Next, we sought to determine some histone modifications that mark active and inactive chromatin in *sns5* and in downstream viral promoter sequences. By using commercially available specific antibodies, we studied the levels of histone H3 acetylation of lysine9 (H3K9ac), which is considered to be a key indicator of an "open" chromatin histone code,[22] and of histone H3 trimethylation of lysine4 (H3K4me3), which is another marker of euchromatin. The levels of H3 dimethylation of lysine9 (H3K9me2) and the recruitment of MeCP2 to methylated CpG dinucleotides were also analyzed as markers of inactive chromatin (**Figure 6b**). Remarkably, in the chromatin of GFP$^+$ clones, *sns5* sequences were highly enriched for H3K9 acetylated nucleosomes when compared to the LTR promoter (12.9 ± 2.1-fold in *sns5* versus 4.7 ± 0.8-fold in LTR). In the GFP$^-$ clones, a lower enrichment of this histone modification was observed (7.1 ± 1.6-fold in *sns5* versus 1.2 ± 0.7-fold in LTR). The histone H3 trimethylation of lysine4 was highly represented in the LTR promoter of positive clones when compared to the inactive clones (13.6 ± 1.6-fold versus 5.3 ± 1-fold). Intriguingly, comparable levels of this marker were found in *sns5* sequences in both the active and inactive clones (approximately eightfold enrichment). In GFP$^-$ clones, the dimethylation of the H3K9 nucleosomes (*sns5* 4.7 ± 1.3-fold; LTR 5.7 ± 0.7-fold) was more relevant, and the recruitment of the CpG methylation is specific to the MeCP2 protein in the insulator (8 ± 0.9-fold) and LTR (2.9 ± 0.8-fold). Conversely, the levels of H3k9me2 and MeCP2 were undetectable for both sequences in the GFP$^+$ clones. Taken together, these results correlate with previous large-scale and genome-wide studies, which showed a high prevalence of H3K4me3 and H3K9ac-modified nucleosomes that were in close proximity to the transcription start sites and promoter regions of actively transcribed genes.[23–25] These results also correlate with other studies that have shown that chromatin insulators provide a center for histone acetylation.[4,26,27] In conclusion, *sns5* is capable of binding transcription factors and hyperacetylated nucleosomes, which are properties that are common to other well-characterized insulator elements.

DISCUSSION

Retroviral-vector integration into heterochromatin regions can result in both the loss of transgene expression (PE) and the process of PEV. The incorporation of chromatin insulators should limit the spread of heterochromatin, which would allow the gene to be expressed in almost all cells irrespective of integration site. When considering vector-mediated gene therapy, this boundary property is useful to limit the number of integrants per cell necessary to achieve therapeutic effects, and to decrease the risk of insertional mutagenesis. The enhancer-blocking function may also have safety implications. It has been shown that RV and LV vectors have preferential intergenic/intragenic integration sites and that vector enhancers can activate cellular genes[28–30] as well as cause insertional mutagenesis events, as reported in some gene therapy clinical trials.[31,32] The chicken-insulator element, cHS4,

has been shown to alleviate PEV when flanking integrated transgenes, even if it is also reported that this element works best in the context of globin genes and may not be as effective at "insulating" nonerythroid genes.[8,33,34] Here, we report the ability of a new chromatin insulator, *sns5*, from the sea urchin *Paracentrotus lividus H2A* early histone locus, to improve retroviral encoded transgene expression in murine erythroid milieu. We found that flanking the recombinant retroviral vector, MGPN2, with the *sns5* element reduced the rate and the severity of positional effects on vector transcription. We first demonstrated that the inclusion of the *sns5* element did not affect vector titer and stability, which are crucial prerequisites for its use with retrovirus vectors. The likelihood of vector expression was investigated in MEL cells by avoiding selection based on transgene expression that can influence the degree of chromosomal PE. By limiting dilution assays, we showed that the *sns5* chromatin insulator, when flanking the transcription unit in a forward orientation, improved the likelihood of transgene expression, as it almost doubled the percentage of cells that expressed the vector. Further analysis that was performed in GFP^+ and GFP^- fractions after transduction and during cell expansion allowed us to demonstrate that the *sns5* element is able to counteract the initial silencing and the extinction of expression. In fact, in the GFP sorted negative fraction, we found multiple copies of the uninsulated provirus per cell as opposed to only 0.1 copies per cell of the *sns5* insulated provirus, which suggested that a significant number of uninsulated provirus copies are silenced immediately. Following the extinction profile of the LTR→GFP expression cassette in GFP-expressing cells, it was evident that further copies of the uninsulated provirus were silenced later, within the first week, as the percentage of GFP-expressing cells dramatically decreased at that time. That the *sns5* element was able to counteract PEV is also evident at the clonal level, as individual clones that contained single integrants showed that the majority of cell progeny expressed the transgene after 30 days of cell expansion, whereas the expression of the uninsulated vector was highly variegated at the same time point. The results obtained in NIH-3T3 cells indicate that *sns5* (S) also increases the probability of transgene expression within cell progeny of individual clones in nonerythroid cells. The boundary activity of *sns5* is confirmed by a direct correlation between the level of GFP expression and VCN. We had previously reported that the 265 bp *sns5* core element, named *sns,* displayed enhancer-blocking activity in both orientations in erythroid milieu.[17] The results reported here indicate that *sns5*, but not *sns,* confers full insulator activity, which suggests that the two properties of enhancer blocking and PE protection map to different sequences within this element, as already demonstrated for other insulators.[35] The boundary activity of *sns5* probably involves structural aspects and/or redundancy of proteins; in fact in the additional 200 bp of *sns5*, we detected by EMSA a second binding site for GATA1 factor. Incoming experiments will elucidate these aspects. To better understand the mechanisms that underlie the protection against PE that are afforded by the *sns5* sea urchin insulator, we analyzed the *in vivo* binding of transcription factors and the epigenetic modification that marks active and inactive chromatin. By ChIP analysis, we showed the specific binding of the erythroid GATA1 and the ubiquitous OCT1 transcription factors and the recruitment of hyperacetylated nucleosomes in the insulator sequences of the transcriptionally active provirus. These findings are in line with previously suggested insulator mechanisms of action. It is known that insulators alter the structure of chromatin by affecting the covalent modifications of histones, which has been demonstrated in both yeast and vertebrate cells. These studies led to a model in which the boundary of the silenced region is determined by a dynamic equilibrium between histone deacetylation activity that originates from the heterochromatin and histone acetylation activity that is centered on the barrier element. The barrier activity of the native cHS4 element is associated with a peak of histone hyperacetylation and other hallmarks of an "open" histone code, which in turn is thought to prevent the invasion of transcriptionally repressive heterochromatin from an adjoining region into the chicken β-globin locus.[26] It has been recently shown that the increase in the likelihood of expression of a reporter gene in a cHS4-insulated γ-retrovirus vector was correlated with an increase in histone H3 acetylation, which peaked at the insulator sequences.[27] Similar results are reported in this study, which has shown a 13-fold enrichment in the level of histone H3K9ac in the *sns5* sequence versus a 4.7-fold enrichment in the downstream LTR promoter sequences. It is also known that transcription factors may contribute to blocking heterochromatin propagation through the recruitment of histone-modifying complexes, such as histone acetylases.[36,37] The erythroid transcription factor GATA1 was reported to be required to establish a tissue-specific histone acetylating pattern by interacting with the acetyl-transferase CBP/P300, which mediates changes in histone acetylation.[38,39] Our findings of colocalization in the insulator sequences of GATA1, other transcription factors, and acetylated nucleosomes have significant implications for the conservation of insulator function in evolutionary distant organisms, which suggests that this new insulator is able to maintain a euchromatin state inside the proviral locus with mechanisms that are common to other characterized insulators. Other epigenetic modifications that are associated with active chromatin, such as H3K4me3, are present in the transcriptionally active provirus, whereas markers of silent chromatin, such as H3K9me2 and the recruitment of MeCP2 to methylated CpG dinucleotides, are undetectable. Further experiments are necessary to better outline the role of these factors in reducing chromosomal PE on integrated vector transcription. The relatively small size of the *sns5* insulator should simplify modifications of vectors and make some constructions possible, where the total length of the insert should be limited. The capability to improve the likelihood of retroviral-vector expression in erythroid cells could be of great interest for its potential application in gene therapy of erythroid disorders, such as β-thalassemia.

MATERIALS AND METHODS

Retrovirus vectors. The retrovirus-vector constructs, which are shown in **Figure 1**, were generated by using the murine stem cell virus vector, MGPN2.[40] This vector contains the viral LTRs, the extended packaging signal from the murine stem cell virus that expresses the *GFP* gene from the viral 5′-LTR promoter (LTR→GFP cassette) and the *NEO* gene from a PGK promoter (PGK→NEO cassette).[41] The recombinant vectors were designed by inserting different DNA fragments into the Not1-restriction site of the multiple cloning site, which were previously inserted into the Nhe I site at the 3′-LTR of the original MGPN2 vector. From this

location, the insulator fragments are copied into the 5′-LTR during provirus integration in order to generate a flanking "double-copy" configuration.[42] To generate *sns* (S) and *sns* (R) vectors, a 265 bp HpaII-Sau3A *sns* fragment, which was from the subclone of pH70 histone DNA,[43] was cloned into the pGEM T-easy plasmid, cut with the Not1-restriction enzyme and inserted into the 3′-LTR in both orientations with respect to virus transcription. The *sns5* (S) and *sns5* (R) vectors were constructed similarly by inserting a Not1-462 bp-*sns5* fragment from the H2A-CAT-*sns5* plasmid.[18] As a control, spacer DNAs of comparable sizes from λ-phage DNA were cloned into the same position, which generated the 270-spacer vector and the 564-spacer vector. The cHS4 vector, which contains the 1,203 bp XbaI fragment from the chicken *β-globin* locus inserted in 5′–3′ orientation at the Nhe I site of the LTR, was kindly provided by D.W. Emery (University of Washington, Seattle, WA). Vector producer clones were generated under G418 selection (800 μg/ml) by transfection with the GP+E86 ecotropic packaging line[44] (ProFection Mammalian Transfection Systems; Promega, Madison, WI). Virus titers were determined by serial dilution and transfer of G418 resistance to naive NIH-3T3 cells from 15 producer clones for each construct. Genomic DNAs were analyzed for intact provirus and VCN by Southern blot. The absence of replication-competent viruses was determined by a standard marker-rescue assay.[45] Producer clones with the highest titers were used for the MEL cell transduction. Vector-containing supernatant was collected from semiconfluent GP+E86 producer clones and passed through a 0.45 μm filter.

Cell lines. The mouse fibroblast cell line NIH3T3, the packaging cell line GP+E86, and the adult-stage murine erythroleukemia cell line MEL585[46] were all maintained in Dulbecco's modified Eagle's medium (Gibco Invitrogen Life Science Technology, Grand Island, NY) supplemented with 10% heat-inactivated fetal bovine serum (Gibco), 2 mmol/l L-glutamine (Gibco), 1 mmol/l sodium pyruvate (Gibco), 0.1 mmol/l nonessential amino acids (Gibco) and the antibiotics penicillin/streptomycin (10 μg/ml) (Gibco) at 37 °C in the presence of 5% CO_2.

MEL cell transduction. A total of 5×10^5 MEL cells were transduced at a modest multiplicity of infection in order to minimize provirus copy number by a 24-hour culture in vector-containing supernatant in the presence of 10 μg/ml polybrene (Sigma Chemical, St Louis, MO). The cells were then washed and plated in the absence of selection at a limiting dilution in 96-well flat-bottomed dishes. Individual clones were expanded and screened for the presence of vector and its expression. To study the extinction profile of the *GFP*-reporter gene, GFP^+ cells were recovered by sorting (EPICS ALTRA HYPERSORT; Beckman Coulter, Miami, FL) 72 hrs after transduction, with the percentage of GFP^+ cells measured twice a week by flow cytometer (FC-500; Beckman Coulter).

DNA analysis. To determine provirus integrity, genomic DNA was isolated by standard methods, and 10 μg was digested with the *Eco*RV restriction enzyme, which cuts once in each viral LTR. Southern-blot analysis was then performed using a P^{32}-labeled 923 bp *PstI* fragment from the *NEO* gene as a probe. The full-length provirus band was 3,591 bp long for the MGPN2 vector, 4,053 bp for the *sns5* (S) and *sns5* (R) vectors, 3,856 bp for the *sns* (S) and *sns* (R) vectors, and 4,791 bp for the cHS4 vector. Using the same probe, the VCN/cell in individual clones was detected by digestion of genomic DNA with *BamHI*, which cuts once in the vector and generates a number of bands corresponding to the integration events. VCN/cell in pools was determined by digesting the DNA with *Eco*RV and *BamHI* and hybridizing the blot simultaneously with the *NEO* probe and with a 369 bp *Eco*RI fragment from the mouse *β-globin* locus *εy* as the other probe. The hybridization signals from the 3.33 kb provirus band and the 3.67 kb endogenous *εy* gene band were quantified with a PhosphoImager (Biorad GS525, Milano, Italy). To detect the presence of integrated vectors, multiplex PCR analysis was carried out using *NEO* gene specific oligos (5′-GCAAATCGGCTGCTCTGATG-3′ and 5′-CTCGCTGGATGCGATGTTTC-3′) and oligos that were specific for mouse *β-globin* locus as a control (5′-TGACTCAGCAAACCCTAGGC-3′ and 5′-CTTTGTCCTACTGCTCTCATG-3′). Two PCR fragments, which were 350 bp from the *NEO* gene and 220 bp from mouse LCR hypersensitive site HS3, were generated. The double-copy configuration of the inserted DNA fragments was detected by PCR analysis using a *GFP* specific oligonucleotide as reverse primer (5′-CTGAACTTGTGGCCGTTTA-3′) for all the constructs and different oligonucleotides as forward primers, which depended on the inserted sequence and orientation (SP3 primer: 5′-GTCGATCATCCGATCTAATAT-3′, for *sns* and *sns5* that are cloned in forward orientation; SP20 primer: 5′-GACAACAACTACAA AACTGGAAT-3′, for *sns* and *sns5* that are cloned in reverse orientation; cHS4 primer: 5′-AACATGCAGGCTCAGACACA-3′). Sequence analysis of the PCR products was carried out to confirm the fidelity of the retrotransciption process (Genetic Analyzer ABI PRISMA 3130 XL; Applied Biosystems, Foster City, CA).

ChIP analysis. MEL cells that were transduced with the *sns5* (S) vector were subcloned and screened for cells carrying transcriptionally active or inactive single-copy vectors. The ChIP assay was carried out as described elsewhere,[47] with minor modifications. Briefly, protein-DNA crosslinking was performed by incubating 3×10^7 cells with 0.4% formaldehyde for 10 minutes at room temperature. Soluble chromatin complex in 1 ml of NICaCl$_2$ buffer (15 mmol/l Tris pH 7.5, 1 mmol/l $CaCl_2$, 60 mmol/l KCl, 0.5 mmol/l DTT, 15 mmol/l NaCl, 300 mmol/l sucrose) was produced by digestion with micrococcal nuclease from *Staphylococcus aureus* (Sigma Life Science, St Louis, MO) (1.5 U/ml) for 7 minutes at 37 °C. Under our conditions, the length of the digested chromatin ranged from 0.2 to 0.6 kb. To reduce the nonspecific background, the samples were incubated with 100 μl of salmon sperm DNA/protein-A sepharose (Sigma Life Science) slurry for 1 hour at 4 °C with agitation. Twenty percent of chromatin that was cleared by centrifugation (input DNA) was withdrawn and processed as the immunoprecipitated chromatin. For each ChIP experiment, 25 μg of the DNA that contained chromatin was incubated in ChIP lysis buffer (1% SDS, 10 mmol/l EDTA, 50 mmol/l Tris-HCl pH 8) overnight at 4 °C with 5 μg of each of the following antibodies anti-GATA1 (Santa Cruz Biotechnology, Santa Cruz, CA), anti-OCT1 (Santa Cruz Biotechnology), anti-MeCP2 (H300; Santa Cruz Biotechnology), anti-acetyl-histone H3 (Lys9) (H3K9ac; Upstate, Charlottesville, VA), anti-thrimetyl-histone H3 (Lys4) (H3K4me3; Upstate), anti-dimethyl-histone H3 (Lys 9) (H3K9me2; Upstate). The same chromatin aliquot was incubated with buffer only and used as negative control (no antibody). Five nanograms of the immunoprecipitated chromatin and of the input DNA, which was quantified using the Qubit fluorometer (Invitrogen Detection Technologies, Eugene, OR), were used for RT-PCR analysis (7,900 HT ABI PRISM; Applied Biosystem), with the following primers specific for *sns5* and LTR sequences: *sns5* forward: 5′-CCGGCAAATCAAGCTAAAGG-3′, *sns5* reverse: 5′-AGGGCCGTTGAGGTGTTG-3′, *sns5* probe: 5′-TGCACT CGCAAACC-3′ LTR forward: 5′-CGCCCTCAGCAGTTTCTAGAGA-3′, LTR reverse: 5′-AAGGACCTGAAATGACCCTGTG-3′; and LTR probe: 5′-CCATCAGATGTTTCCAGGGT-3′.

ACKNOWLEDGMENTS

We thank David W. Emery (University of Washington, Seattle, WA) for kindly providing the cHS4 vector, Barbara Spina (V. Cervello Hospital, Palermo, Italy) for technical assistance, Filippo Leto (V. Cervello Hospital, Palermo, Italy) for sequence analysis, Daniel R. Hollyman (MSKCC, New York, NY) for helpful discussion. This work was supported by Fondazione Franco e Piera Cutino, Palermo Italy; Telethon Foundation grant no. GGP08221; Fondazione "Monte Paschi di Siena," Rome Italy; "Stem Cell Project" of Fondazione Roma, Rome, Italy. This work was performed at the Unità di Ricerca P. Cutino, Ematologia II, A.O. "V. Cervello," Palermo, Italy.

REFERENCES

1. Pannell, D and Ellis, J (2001). Silencing of gene expression: implications for design of retrovirus vectors. *Rev Med Virol* **11**: 205–217.
2. Karpen, GH (1994). Position-effect variegation and the new biology of heterochromatin. *Curr Opin Genet Dev* **4**: 281–291.
3. Gaszner, M and Felsenfeld G (2006). Insulators: exploiting transcriptional and epigenetic mechanism. *Nat Rew Genet* **7**: 703–713.
4. West, AG, Gaszner, M and Felsenfeld, G (2002). Insulators: many functions, many mechanisms. *Genes Dev* **16**: 271–288.
5. Geyer, PK and Clark, I (2002). Protecting against promiscuity: the regulatory role of insulators. *Cell Mol Life Sci* **59**: 2112–2127.
6. Bell, AC and Felsenfeld, G (1999). Stopped at the border: boundaries and insulators. *Curr Opin Genet Dev* **9**: 191–198.
7. Bell, AC, West, AG and Felsenfeld, G (2001). Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* **291**: 447–450.
8. Rivella, S, Callegari, JA, May, C, Tan, CW and Sadelain, M (2000). The cHS4 insulator increases the probability of retroviral expression at random chromosomal integration sites. *J Virol* **74**: 4679–4687.
9. Emery, DW, Yannaky, E, Tubb, J and Stamatoyannopoulos G (2000). A chromatin insulator protects retrovirus vectors from position effects. *Proc Natl Acad Sci USA* **97**: 9150–9155.
10. Emery, DW, Yannaki, E, Tubb, J, Nishino, T, Li, Q and Stamatoyannopoulos, G (2002). Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing *in vivo*. *Blood* **100**: 2012–2019.
11. Yannaki, E, Tubb, J, Aker, M, Stamatoyannopoulos, G and Emery, DW (2002). Topological constraints governing the use of the chicken HS4 chromatin insulator in oncoretrovirus vectors. *Mol Ther* **5**: 589–598.
12. Arumugam, PI, Scholes, J, Perelman, N, Xia, P, Yee, JK and Malik P (2007). Improved human beta-globin expression from self-inactivating lentiviral vectors carrying the chicken hypersensitive site-4 (cHS4) insulator element. *Mol Ther* **15**: 1863–1871.
13. Ryu, BY, Persons, DA, Evans-Galea, MV, Gray, JT and Nienhuis, AW (2007). A chromatin insulator blocks interactions between globin regulatory elements and cellular promoters in erythroid cells. *Blood Cells Mol Dis* **39**: 221–228.
14. Palla, F, Melfi, R, Anello, L, Di Bernardo, M and Spinelli G (1997). Enhancer blocking activity located near the 3′ end of the sea urchin early H2A histone gene. *Proc Natl Acad Sci USA* **94**: 2272–2277.
15. Melfi, R, Palla, F, Di Simone, P, Alessandro, C, Calì, L, Anello, L *et al.* (2000). Functional characterization of the sea urhin early histone gene cluster reveals insulator properties and three essential cis-acting sequences. *J Mol Biol* **304**: 753–763.
16. Di Simone, P, Di Leonardo, A, Costanzo, G, Melfi, R and Spinelli, G (2001). The sea urchin sns insulator blocks CMV enhancer following integration in human cells. *Biochem Biophys Res Commun* **284**: 987–992.
17. Acuto, S, Di Marzo, R, Calzolari, R, Baiamonte, E, Maggio, A and Spinelli, G (2005). Functional characterization of the sea urchin sns chromatin insulator in erythroid cells. *Blood Cells Mol Dis* **35**: 339–344.
18. Di Caro, D, Melfi, R, Alessandro, C, Serio, G, Di Caro, V, Cavalieri, V *et al.* (2004). Down-regulation of early sea urchin histone H2A gene relies on cis regulative sequences located in the 5′ and 3′ regions and including the enhancer blocker sns. *J Mol Biol* **342**: 1367–1377.
19. Hawley, RG, Lieu, FH, Fong, AZ and Hawley, TS (1994). Versatile retroviral vectors for potential use in gene therapy. *Gene Ther* **1**: 136–138.
20. Aker, M, Tubb, J, Miller, DG, Stamatoiannopoulos, G and Emery, DW (2006). Integration bias of gamma-retrovirus vectors following transduction and growth of primary mouse hematopoitic progenitors cells with and without selection. *Mol Ther* **4**: 226–235.
21. Karpen, GH (1994). Position-effect variegation and the new biology of heterochromatin. *Curr Opin Genet Dev* **4**: 281–291.
22. Margueron, R, Trojer, P and Reinberg, D (2005). The key to development: interpreting the histone code. *Curr Opin Genet Dev* **15**: 163–176.
23. Bernstein, BE, Kamal, M, Lindblad-Toh, K, Bekiranov, S, Bailey, DK, Huebert, DJ *et al.* (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**: 169–181.
24. Kim, TH, Barrera, LO, Zheng, M, Qu, C, Singer, MA, Richmond, TA *et al.* (2005). A high-resolution map of active promoters in the human genome. *Nature* **436**: 876–880.
25. Roh, TY, Cuddapah, S, Cui, K and Zhao, K (2006). The genomic landscape of histone modifications in human T cells. *Proc Natl Acad Sci USA* **103**: 15782–15787.
26. Litt, MD, Simpson, M, Gaszner, M, Allis, CD and Felsenfeld, G (2001). Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293**: 2453–2455.
27. Li, C and Emery DW (2008). The cHS4 chromatin insulator reduces gammaretroviral vector silencing by epigenetic modification of integrated provirus. *Gene Therapy* **15**: 49–53.
28. Schroder, AR, Shinn, P, Chen, H, Berry, C, Ecker, JR and Bushman, F (2002). HIV-1 integration in the human genome favors active genes and local hotspots. *Cell* **110**: 521–529.
29. Wu, X, Li, Y, Crise, B and Burgess, SM (2003). Transcription start regions in the human genome are favored targets for MLV integration. *Science* **300**: 1749–1751.
30. von Kalle, C, Baum, C and Williams, DA (2004). Lenti in red: progress in gene therapy for human hemoglobinopathies. *J Clin Invest* **114**: 889–891.
31. Hacein-Bey-Abina, S, Von Kalle, C, Schmidt, M, McCormack, MP, Wulffraat, N, Leboulch, P *et al.* (2003). LMO2- associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415–419.
32. Ott, MG, Schmidt, M, Schwarzwaelder, K, Stein, S, Siler, U, Koehl, U *et al.* (2006). Correction of X-linked chronic granulomatous disease by gene therapy, augmented by insertional activation of MDS1-EVI1, PRDM16 or SETBP1. *Nat Med* **12**: 401–409.
33. Walters, MC, Fiering, S, Bouhassira, EE, Scalzo, D, Goeke, S, Magis, W *et al.* (1999). The chicken beta-globin 5′HS4 boundary element blocks enhancer-mediated suppression of silencing. *Mol Cell Biol* **19**: 3714–3726.
34. Ramezani, A, Hawley, TS and Hawley, RG (2003). Performance- and safety-enhanced lentiviral vectors containing the human interferon-beta scaffold attachment region and the chicken beta-globin insulator. *Blood* **1101**: 4717–4724.
35. Recillas-Targa, F, Pikaart, MJ, Burgess-Beusse, B, Bell, AC, Litt, MD, West, AG *et al.* (2002). Position-effect protection and enhancer blocking by the chicken β-globin insulator are separable activities. *Proc Natl Acad Sci USA* **99**: 6883–6888.
36. Kuhn, EJ and Geyer PK (2003). Genomic insulators: connecting properties to mechanism. *Curr Opin in Cell Biol* **15**: 259–265.
37. Huang, S, Li, X, Yusufzai, TM, Qiu, Y and Felsenfeld, G (2007). USF1 recruits histone modification complexes and is critical for maintenance of a chromatin barrier. *Mol Cell Biol* **27**: 7991–8002.
38. Letting, DL, Rakowski, C, Weiss, MJ and Blobel, GA (2003). Formation of a tissue-specific histone acetylation pattern by the hematopoietic transcription factor GATA-1. *Mol Cell Biol* **23**: 1334–1340.
39. Blobel GA (2000). CREB-binding protein and p300: molecular integrators of hematopoietic transcription. *Blood* **95**: 745–755.
40. Cheng, L, Du, C, Murray, D, Tong, X, Zhang, YA, Chen, bp *et al.* (1997). A GFP reporter system to assess gene transfer and expression in human hematopoietic progenitor cells. *Gene Ther* **4**: 1013–1022.
41. Hawley, RG, Lieu, FH, Fong, AZ and Hawley, TS (1994). Versatile retroviral vectors for potential use in gene therapy. *Gene Ther* **1**: 136–138.
42. Hantzopoulos, PA, Sullenger, BA, Unger, G and Gilboa E (1989). Improved gene expression upon transfer of the adenosine deaminase minigene outside the transcriptional unit of a retroviral vector. *Proc Natl Acad Sci USA* **6**: 2895–2902.
43. Spinelli, G, Gianguzza, F, Casano, C, Acierno, P and Burckhardt, J (1979). Evidences of two different sets of histone genes active during embryogenesis of the sea urchin Paracentrotus lividus. *Nucleic Acids Res* **6**: 545–560.
44. Markowitz, D, Goff, S and Bank, A (1988). Construction and use of a safe and efficient amphotropic packaging cell line. *Virology* **167**: 400–406.
45. Miller, AD and Rosman, GJ (1989). Improved retroviral vectors for gene transfer and ezxpression. *BioTechniques* **7**: 980–890.
46. Enver, T, Brice, M, Karlinsey, J, Stamatoyannopoulos, G and Papayannopoulou, T (1991). Developmental regulation of fetal to adult globin gene switching in human fetal erythroid x mouse erythroleukemia cell hybrids. *Dev Biol* **148**: 129–137.
47. Di Caro, V, Cavalieri, V, Melfi, R and Spinelli, G (2007). Constitutive promoter occupancy by the MBF-1 activator and chromatin modification of the developmental regulated sea urchin α-H2A histone gene. *J Mol Biol* **365**: 1285–1297.