Published in final edited form as:
Blood. 2005 April 15; 105(8): 3346–3352. doi:10.1182/blood-2004-03-0987.

Differences of globin transgene expression in stably transfected cell lines and transgenic mice

Qiliang Li, **David W. Emery**, **Hemei Han**, **Jin Sun**, **Man Yu**, and **George Stamatoyannopoulos**
Department of Medicine, Division of Medical Genetics, University of Washington, Seattle

Abstract

Previous studies demonstrated that DNase I hypersensitive site −40 (HS-40) of the α-globin locus is capable of greatly enhancing expression of a hybrid β/γ-globin transcriptional unit in plasmid-transfected murine erythroleukemia (MEL) cells. However, as reported here, this same γ-globin gene expression cassette was only transcribed at trace amounts in erythroid cells of transgenic mice. This lack of expression was not directly attributable to the β/γ-globin transcriptional unit, since this same unit linked to a composite β-globin locus control region was expressed at high levels in transgenic mice. This lack of expression was also not directly attributable to chromosomal position effects, since addition of chromatin insulators failed to increase the frequency of expression. DNase I hypersensitivity and chromatin immunoprecipitation assays demonstrated that the lack of expression was correlated with a closed chromatin structure. We hypothesize that transgenes undergo dynamic changes in chromatin conformation following chromosomal integration and that the discrepant results reported here can be attributed to the relatively high level of chromatin remodeling that occurs in the transgenic mouse model, coupled with the relative inability of the HS-40 element to maintain an open chromatin state under such conditions.

Introduction

Several studies suggest that DNase hypersensitive site −40 (HS-40) of the human α-globin locus may prove useful for enhancing expression of recombinant virus vectors encoding human β-globin or γ-globin designed to treat patients with β-chain hemoglobinopathies. This element has been shown to elevate expression of linked α-globin genes in a variety of settings, including cells lines and transgenic mice,[1–4] as well as linked γ-globin and β-globin genes in cell lines.[5,6] The HS-40 element can also be included in recombinant oncoretrovirus vectors for human γ-globin without diminishing vector titers and genetic stability,[5–7] a significant problem with similar elements from the β-globin locus control region (LCR).[8] In previous studies we reported development of an optimized expression cassette for human γ-globin consisting of the HS-40 enhancer linked to a truncated β-globin promoter (βpr) and γ-globin coding segment containing a partial deletion of the second intron.[6] Using a stable plasmid transfection assay in a murine erythroleukemia (MEL) cell line, we found that this cassette expressed γ-globin at 166% of endogenous mouse α-globin per copy in 6 pools and 105% in 9 clones.[6] Following minor modifications, this cassette was introduced into a genetically stable, high-titer oncoretrovirus vector, which was capable of expressing γ-globin at up to 66% per copy of α-globin in MEL studies.[7,9] To investigate how well the γ-globin gene of this optimized expression cassette would be expressed in erythroid cells we produced transgenic mice. In contrast with the high-level expression in stably transfected MEL cells, this γ-globin gene was only transcribed in a trace amount in the erythroid cells of transgenic mice. This unexpected result prompted us to perform a systematic study of expression potential of the γ-globin gene

Reprints: Qiliang Li, Medical Genetics, Box 357720, University of Washington, Seattle, WA 98195; li111640@u.washington.edu.

expression cassette in transgenic mice. These studies suggest that it is a failure of the HS-40 element, rather than elements of the hybrid βpr/γ gene transcriptional unit, that accounts for the gene silencing observed in vivo. Further, these studies suggest that this failure reflects an inherent property of the HS-40 element rather than the influence of silencing chromosomal position effects, since inclusions of chromatin insulators fail to rescue expression. We hypothesize that transgenes undergo dynamic changes in chromatin conformation after integration into target cell chromosomes and that differences in transgene expression in transfected MEL cells and transgenic mice can be explained by differences in the kinetics or degree of chromatin remodeling or both which occur in these 2 settings, coupled with the inability of the HS-40 element to maintain an open chromatin structure during the high level of chromatin remodeling that occurs during embryogenesis.

Materials and methods

Transgene constructs

All constructs were generated in the plasmid vector pBluescript (Stratagene, La Jolla, CA) with the use of standard molecular cloning procedures and are diagramed in Figure 1. As previously described, construct HS-40/βpr/γ gene contains the HS-40 element from the human α-globin locus as a 356–base pair (bp) *Taq*I/*Xmn*I fragment, the human β-globin promoter as a 178-bp *Rsa*I/*Nco*I fragment spanning from the CACCC box to the ATG starting codon, and the human ^{A}γ-globin coding gene as a 1185-bp *Nco*I/*Hin*dIII fragment with a 714-bp *Xho*I/*Hpa*I deletion within the second intron.6 For construct HS-40(mut)/βpr/γ gene a G-to-T transversion was introduced into the 3′ NF-E2 motif of HS-40 with the use of the QuickChange site-directed mutagenesis kit (Stratagene) and was verified by DNA sequencing. For construct μLCR/βpr/γ gene the HS-40 element was replaced with a previously described 2.5-kb composite of the β-globin LCR (μLCR) that comprises 5′ HS1 (GenBank humhbb 13062–13769), HS2 (8486–9218), HS3 (4608–5172), and HS4 (1182–1702).[10] For construct HS-40/βpr/γ gene/hHS5 a previously described 3.2-kilobase (kb) *Bgl*II/*Bam*HI fragment containing HS5 of the human β-globin LCR (accession number L22754) was added 3′ of the γ-globin coding segment.[11] For construct cHS4/HS-40/βpr/γ gene a previously described 1.2-kb *Xba*I fragment containing HS4 of the chicken β-globin LCR (accession number U78775) was added 5′ of the HS-40 element.[12]

Transgenic mice

The construct fragments were released from plasmid sequences by restriction digestion and were purified by agarose gel electrophoresis. The purified fragments were injected into fertilized eggs from B6/C3F1 donors, which were then transferred to pseudopregnant B6/D2F1 foster recipients. Founder animals were identified by hybridizing Southern slot blots with ^{32}P-labeled probes for construct-specific sequences. F1 progeny were obtained by breeding founder animals with nontransgenic B6/D2F1 mates and were screened for correct integration and to exclude the presence of mosaicism in the founders.

Transgene structural analysis

DNA was isolated by standard procedures and digested with *Nco*I, which cuts once within each transgene. DNA (10 μg) was loaded onto an 0.8% agarose gel, and DNA fragments were resolved by electrophoresis. Southern blot hybridization was performed with ^{32}P-labeled probes and quantified on a PhosphorImager (Molecular Dynamics, Sunnyvale, CA). For copy number measurements the blots were hybridized with a mixture of 2 probes for the recombinant construct and 1 probe for the murine α-globin gene as a loading control. Copy numbers were determined by comparing the signals from a given transgenic line with those from human genomic DNA.

RNA analysis

Total RNA was prepared from peripheral blood of 3- to 5-week-old transgenic mice with the use of guanidine thiocyanate-acid–phenol. Human γ-globin transgene transcripts were analyzed by RNase protection as previously described by using the following probes: pT7Mα (128) linearized with *Hin*dIII to give a 128-bp protected fragment within exon 1 of the mouse α-globin gene and pT7^{A}γ (170) linearized with *Bst*EII to give a 170-bp protected fragment within exon 2 of the human ^{A}γ-globin gene.[13] A total of 50 ng RNA was hybridized overnight at 48°C with 10^6 counts per minute (cpm) of each radiolabeled probe. A pilot experiment confirmed that probe was in excess under these conditions. After digestion with RNase A and T1, the protected fragments were separated on 6% polyacrylamide-8M urea gels, and autoradiography was performed without intensifying screens. Signal intensities were quantified by PhosphorImager (Molecular Dynamics). Expression levels of γ-globin were calculated as a percentage per copy of mouse α-globin by using the following formula: human γ-globin mRNA/γ-globin copy number/murine α-globin mRNA/4. To minimize experimental error, multiple measurements were performed on individual animals.

DNase I hypersensitivity analysis

Nuclei were prepared from the spleens of cHS4/HS-40/βpr/γ gene transgenic mice as previously described[14] and subject to limited DNase I treatment. DNA was then purified, digested with *Eco*RI, separated by electrophoresis, blotted, and hybridized with a probe for the 3′ end of the γ-globin gene.

Chromatin immunoprecipitation analysis

These assays were performed as described with minor modifications.[14] In short, splenocytes from cHS4/HS-40/βpr/γ gene transgenic mice were collected and cross-linked with formaldehyde. After sonication, soluble chromatin was immunoprecipitated with rabbit polyclonal antibodies against histone H3 acetylated at lysines 9 and 14 (catalog 06-599) or histone H4 acetylated at lysines 5, 8, 12, and 16 (catalog 06-866) purchased from Upstate Biotechnology (Lake Placid, NY). The level of transgene sequences in the precipitated fraction was then determined by real-time polymerase chain reaction (PCR). Controls included the mouse *Aire* gene, which is not expressed in spleen, and the mouse β^{maj}, which is expressed at high levels in spleen. Primers used in PCR include the following: αHS-40, CTGCTGATTACAACCTCTGGTG and TCAGCTCCAGATGAAGAACGTA; β-globin promoter, ACGGCTGTCATCACTTAGACCT and GAACACAGTTGTGTCAGAAGCA; γ-globin exon 1, GCCATGGGTCATTTCACAG and TGGTCACCAGAGCCTACCTT; cHS4, AAGCCCCCAGGGATGTAAT and CGAGAAGCGTTCAGAGGAAAG; mouse *Aire* gene, CACACAGGCAATTTGTCCAC and AGAGGAAGCGGGAAAGTTGT; and mouse β^{maj} gene, GTCATCACCGAAGCCTGATT and TGTCTGTTTCTGGGGTTGTG.

Immunofluorescence staining and flow cytometry

As previously described, peripheral blood cells were suspended in 1 mL Hanks buffered saline solution with 4% formaldehyde and incubated at room temperature for 30 minutes.9,[15] The cells were then permeabilized by serial washes in cold acetone, washed once with cold Hanks buffered saline solution (HBSS)–2% fetal bovine serum (FBS), and stained with a phycoerythrin-conjugated anti–hemoglobin F (HbF) monoclonal antibody (BD Biosciences, San Diego, CA) for 30 minutes on ice. The cells were washed again with HBSS-2% FBS and analyzed by flow cytometry on a FACScan (BD Biosciences) using CellQuest software (BD Biosciences).

Results

Expression of original cassette

Our previous studies revealed that an optimized expression cassette for human γ-globin could promote high-level expression in stably transfected MEL cells.[6] To assess whether this γ-globin expression cassette has similar potential to promote high-level expression in erythroid cells, we sought to analyze expression of this cassette in transgenic mice. As diagramed in Figure 1, this original construct (designated HS-40/βpr/γ gene) consists of the HS-40 regulatory element from the human α-globin locus linked to a truncated β-globin promoter and γ-globin coding segment containing a partial deletion of the second intron. Eight founders were identified by slot blot hybridization of mouse tail DNA by using a γ-globin gene probe. However, RNase-protection analysis of peripheral blood from these founders revealed barely detectable levels of γ-globin (data not shown). To exclude the possibility that the low expression was due to mosaicism of the transgene in the founder mice, F1 transgenic mouse lines were established. The structure of the transgene in the 4 lines was analyzed by Southern blot hybridization. Genomic DNA was digested with *Nco*I, which cuts once within the transgene. As seen in the top panel of Figure 2 and in Table 1, a 1.7-kb fragment was detected in all lines with the use of an HS-40 probe, indicating that the transgene was arranged in head-to-tail repeats. In addition to the repeat band, 1 to 4 other bands, which represent end fragments of the transgene, were detected in each line, suggesting that the transgene integrated in a limited number of sites. The copy numbers of the transgene ranged from 3 to 26. Expression of the γ-globin transgene in peripheral blood was analyzed by RNase protection. As seen in the bottom panel of Figure 2 and in Table 1, γ-globin mRNA levels were very low in the adult blood of all 4 lines, averaging less than 0.1% of endogenous murine α-globin mRNA. The γ-globin gene was also poorly expressed in the 12-day yolk sacs and fetal livers of transgenic mice (data not shown). These results sharply contrast with the expression studies that used this same cassette in stably transfected MEL cells and suggest an even more profound level of gene inactivation than that observed in oncoretrovirus vector transduction studies with an uninsulated version of this cassette following bone marrow transduction and transplantation in mice.[6,9]

Expression following mutation of HS-40

It has been reported that a G-to-T transversion in the NF-E2 motif of the αHS-40 enhancer is able to derepress expression of a reporter gene transcribed by the ξ-globin gene promoter in the adult blood of transgenic mice.[16] This observation prompted us to test whether introduction of the same mutation could activate expression of our hybrid β(pr)/γ gene transcription unit. As diagramed in Figure 1, this G-to-T transversion was introduced into the HS-40 enhancer of our original γ-globin expression cassette to generate construct HS-40(mut)/βpr/γ gene, and this construct was used to generate 2 transgenic mouse lines. As seen in the left panel of Figure 3 and Table 1, structural analyses by Southern blot hybridization showed that the transgene was integrated as head-to-tail repeats in 1 or 2 sites. As seen in the right panel of Figure 3 and Table 1, RNase protection analysis of adult peripheral blood indicated that the γ-globin transgene was still expressed at an extremely low level, ranging from 0.01% to 0.11% of murine α mRNA per copy. These results demonstrate that modifying the NF-E2 motif of the HS-40 element is not sufficient to activate expression of the hybrid β(pr)/γ gene transcription unit and by inference that the mechanism of this transcriptional inactivity is distinct from mechanism by which adult ζ-globin gene expression is silenced.

Expression following replacement of HS-40 by a μLCR

To determine whether the failure of the original HS-40/βpr/γ gene cassette to activate in transgenic mice could be attributed to some aspect of the hybrid β (pr)/γ gene transcription unit (such as the juxtaposition of the β-globin promoter and the γ-globin coding gene or the partial deletion within the second intron of the γ-globin coding gene), we sought to test this

transcription cassette in a setting independent of the HS-40 element. Several studies have demonstrated that linking intact genes for either γ-globin or β-globin to a recombinant version of the β-globin LCR, termed a "μLCR," results in high levels of expression in blood of adult mice.[17] As diagramed in Figure 1, the HS-40 regulatory element was replaced with a 2.5-kb μLCR to generate construct μLCR/βpr/γ gene. As summarized in Table 1, this construct was used to generate 4 lines containing transgene copies ranging from 2 to 13 (Southern data not shown). As seen in Figure 4 and Table 1, RNase protection analysis of adult peripheral blood indicated that, when linked to the μLCR, per-copy expression of the hybrid β(pr)/γ gene transcription unit ranged from 2.1% to 45.6% of murine α mRNA. These results indicate that the hybrid β(pr)/γ gene transcription unit, although sensitive to chromosomal position effects, is fully capable of high-level expression in transgenic mice and that neither the hybrid promoter-gene combination or the internal deletion of γ-globin intron 2 can, by themselves, explain the lack of expression when this cassette is linked to the HS-40 element.

Expression with chromatin insulators

To assess the role of chromosomal position effects on expression of the original HS-40/βpr/γ gene cassette in the context of transgenic mice, we added to this cassette either the HS5 chromatin insulator from the human β-globin LCR[11] or the HS4 chromatin insulator from the chicken β-globin LCR (termed cHS4).[9,12] As diagramed in Figure 1, the first construct, designated HS-40/βpr/γ gene/hHS5, contained a 3.2-kb fragment encompassing the human HS5 insulator element located at the 3′ end. Previous studies showed that this fragment is able to protect a position-sensitive γ-globin gene from position effects in transgenic mice.[11] The top panel of Figure 5 and Table 1 show the results of the structural analyses by Southern blot hybridization of the 5 transgenic lines generated with this construct. By digesting with *Nco*I, which cuts once within the construct, it was again possible to demonstrate that the vast majority of the transgenes were integrated in head-to-tail repeats in all of these lines. In such arrangement each copy of the HS-40/βpr/γ gene expression cassette was bracketed by the hHS5 chromatin insulator except for the first upstream copy, effectively protecting most copies of this cassette from the silencing effects of flanking chromatin. As seen in the bottom panel of Figure 5 and accompanying summary, this construct was still expressed at a very low level, ranging from 0.02% to 3.6% of murine α-globin mRNA per copy. Although the mean value of γ-globin expression was 20-fold higher in these transgenic lines than in the 4 lines containing the γ-globin expression cassette without the hHS5 insulator, this difference was not statistically significant and was much lower then the level of expression seen when the HS-40/βpr/γ gene cassette was linked to the μLCR. This level of expression was also much lower than that seen with the HS-40/βpr/γ gene cassette in plasmid-transfected MEL cells or retrovirus vector-transduced mouse bone marrow.

As diagramed in Figure 1, the second construct, designated cHS4/HS-40/βpr/γ gene, contained a 1.2-kb fragment encompassing the cHS4 insulator element located at the 5′ end. Previous studies showed that this fragment is able to protect a number of position-sensitive genes from silencing position effects in transfected cell lines, retrovirus vector-transduced bone marrow, and transgenic mice and is the most well-characterized mammalian insulator to date.[9,12,18] The top panel of Figure 6 and Table 1 show the results of the structural analyses by Southern blot hybridization of the 4 transgenic lines generated with this construct. This analysis demonstrated that the vast majority of the transgenes were integrated in head-to-tail repeats in all of these lines, so that most copies of the HS-40/βpr/γ gene expression cassette were bracketed with the cHS4 chromatin insulator. As seen in the bottom panel of Figure 5 and in Table 1, this construct was also expressed at a very low level, ranging from 0.01% to 0.12% of murine α-globin mRNA per copy. Combined with the studies described earlier in this section with the hHS5 insulator, these results indicate that the lack of expression for the HS-40/βpr/γ gene expression cassette in transgenic mice cannot be contributed solely on a sensitivity to

chromosomal position effects but must instead reflect some innate property of the expression cassette itself.

The insulated transgene chromatin is in closed conformation

To determine whether the lack of transgene expression was associated with a specific chromatin structure, the cHS4/HS-40/βpr/γ gene construct was analyzed for DNase I hypersensitivity in transgenic splenocytes. As seen in Figure 7A (right panel), a faint subband from DNase I digestion could be detected at a position corresponding to the human β gene promoter. However, the sensitivity of this subband was 8- to 10-fold less than that of the endogenous β^{maj} gene promoter (Figure 7A, left panel). The chromatin structure of this construct in transgenic splenocytes was also assessed by ChIP with an antibody for acetylated H3, a hallmark of open chromatin. As seen in Figure 7B, all portions of the transgene analyzed exhibited very low levels of acetylated H3, ranging from 0.5- to 0.8-fold of the level associated with the endogenous mouse *Aire* gene used as a transcriptionally silent control. This is in contrast to the endogenous β^{maj} loci, which exhibited acetylated H3 at a 3.3-fold higher level then the control *Aire* gene. Similar results were obtained when the antibody against acetylated histone H4 was used (data not shown). Collectively, these results indicate that the chromatin of the transgene is in a closed state, even when flanked by the cHS4 chromatin insulator.

Expression is heterocellular

Finally, we sought to determine whether the HS-40/βpr/γ gene constructs were expressed in a pancellular pattern, indicating that the transgene was always transcriptionally active but at a very low level, or in a heterocellular pattern, indicating that the transgene was sometimes active and sometimes completely silent. For this purpose we analyzed RBCs from F3 progeny of transgenic line R containing the cHS4/HS-40/βpr/γ gene transgene by immunofluorescent staining and flow cytometry. As seen in Figure 7C, all 6 mice analyzed exhibited a heterocellular pattern of expression, with the fraction of RBCs expressing γ-globin ranging from 1.6% to 6.1%. These results indicate that the very-low overall levels of γ-globin expression observed in these animals arose in large part from a complete failure of the transgene to become transcriptionally active in most cases during erythropoiesis. In addition, comparison of the RNase protection and flow cytometric analysis suggest that the transgene was only expressed at about 1% to 4% of murine α-globin mRNA per copy in the rare cells in which the cassette was transcriptionally active.

Discussion

Several previous studies led us to believe that the HS-40/βpr/γ gene cassette would be expressed in transgenic mice. HS-40 has been identified as the major tissue-specific regulator of the α-globin gene locus.[1] Naturally occurring deletions of this element result in α-thalassemia,[19] while addition of this element to a minimal coding cassette for α-globin results in high levels of tissue-specific expression in transgenic mice and both transiently and stably transfected MEL cells.[2–4,20] Our own previous studies demonstrated that this element could efficiently enhance expression of the hybrid βpr/γ gene transcription unit in MEL cells[6,7] and, to a lesser degree, mice that received transplants with oncoretrovirus vector-transduced marrow.[9] However, it is clear that the HS-40 element does not possess the robust chromatin opening activity of other regulatory elements such as the β-globin LCR.[21] The α-globin gene cluster exists within a naturally occurring open chromosomal environment that is independent of the HS-40 element.[1,21] As a result, the HS-40 element need not create or maintain chromatin in an open, transcriptionally active state but merely enhance transcription from preformed transcriptional complexes at the α-globin gene promoters.[1] Indeed, naturally occurring deletions, which simply result in the juxtaposition of the HS-40 element next to silent chromatin, can result in α-thalassemia.[22] Even the studies with transgenic mice mentioned

here[2–4,20] revealed that the HS-40 element cannot direct copy number dependent or developmentally stable expression in vivo. Further, our own previous studies and those of others found that the HS-40 element frequently failed to activate expression of recombinant γ-globin cassettes following oncoretrovirus vector transduction of mouse bone marrow,[7,23,24] even when the vector was flanked with chromatin insulators.[9]

Nevertheless, we were still surprised by the degree to which the HS-40 element failed to activate the βpr/γ gene transcription unit in transgenic mice. This outcome cannot be attributed solely to the recombinant βpr/γ gene transcription unit, since this same unit, when linked to a μLCR, was expressed at relatively high levels. Likewise, this outcome cannot be attributed solely to the effects of the chromatin environment surrounding the sites of integration, since inclusion of either the hHS5 or cHS4 chromatin insulators failed to increase the level of expression. Rather, the failure of the HS-40/βpr/γ gene cassette to become transcriptionally active in transgenic mice must reflect either some inherent property of the HS-40 element itself or a unique interaction between this element and the βpr/γ gene transcription unit.

There are several possible explanations for the notable discrepancy in the pattern and level of expression for the HS-40/βpr/γ gene cassette in transfected MEL cells, mice that received transplants with oncoretrovirus vector-transduced bone marrow, and transgenic mice. It is unlikely that this discrepancy is due to differences in the milieu of transcription factors present during the later stages of in vivo erythropoiesis and MEL cell differentiation, since both environments are capable of supporting expression of β-globin, γ-globin, and α-globin. It is possible that this discrepancy is somehow related to differences in the amount of time between gene transfer and expression analysis, which typically ranged from 2 to 4 weeks for the MEL cell studies, 1 to 6 months for the bone marrow transduction studies, and 2 generations for the transgenic animal studies. However, extended MEL cell culture studies and secondary bone marrow transplantation studies revealed a stable pattern of expression, indicating that time alone cannot provide a complete explanation. It is also possible that the discrepancy between the in vitro and in vivo results may reflect the fact that plasmid-transfected and oncoretrovirus vector-transduced MEL cells, clones, and pools were isolated according to expression of a selectable drug-resistance marker, while no such selection was exerted in the bone marrow transduction and transgenic mouse studies. In theory, this approach could bias the repertoire of integration events, effectively enriching for integration events that occurred in transcriptionally permissive sites and excluding integration events that failed to support expression of the transferred genes. Indeed, the studies with oncoretrovirus vectors indicate that expression of the HS-40/βpr/γ gene cassette can be highly sensitive to the silencing effects of chromatin surrounding the sites of integration. However, both the previous studies with oncoretrovirus vectors and the transgenic mouse studies reported here clearly demonstrate that inactivity of the HS-40/βpr/γ gene cassette cannot be blamed solely on silencing chromosomal position effects, since this inactivity can, at best, be only partially corrected through the use of chromatin insulators.

Finally, it is possible that the discrepancy between the results in transfected MEL cells, mice that received transplants with oncoretrovirus vector-transduced bone marrow, and transgenic mice may reflect differences in the amount of differentiation and chromatin remodeling that occurs in these different systems. MEL cells represent a nearly fully differentiated state of the erythroid lineage and, thus, are likely to have a relatively stable global chromatin structure. In contrast, the retrovirus vector studies involved gene transfer into primitive progenitor or even more primitive hematopoietic stem cells.[9,24] Several lines of evidence indicate that these more primitive cell populations undergo significant changes in global patterns of chromatin structure, involving active chromatin remodeling.[25,26] Indeed, the high rate of oncoretrovirus vector silencing in the context of bone marrow transduction and transplantation has been attributed in large part to the high degree of chromatin remodeling that takes place in

hematopoietic stem cells.[26,27] Even more profound changes in differentiation and global patterns of chromatin structure occur during germ line transmission and embryogenesis that can effect expression of introduced genes.[26,28,29]

Under this scenario, the differences in the various patterns of expression could be explained as follows. In the case of MEL cells, the transferred cassette would immediately find itself in a transcriptionally supportive cellular erythroid environment, allowing for the early establishment of immature, transcriptionally open chromatin structure.[30] Without a high degree of chromatin remodeling, this transcriptionally active state could readily become fixed, leading to a stable pattern of cassette expression. In the case of gene transfer into repopulating bone marrow, the target cells need to undergo several rounds of division and differentiation before reaching a developmental stage that provides the appropriate cellular erythroid environment to support active expression of the transferred cassette. During this period, the lack of transcriptional activity and a relatively high level of global chromatin remodeling can readily lead to the formation of condensed, transcriptionally silent heterochromatin on the transferred DNA. Without the ability to open such condensed heterochromatin, the HS-40 element would be powerless to activate transcription of the linked expression cassette. In the case of transgenic mice, the rounds of cell division and differentiation, as well as degree of global remodeling, prior to the generation of an erythroid environment are even greater, leading to an even higher frequency of transcriptional inactivity.

Also under this scenario, the intermediate pattern of expression observed in the case of oncoretrovirus vector–transduced bone marrow could either be explained by the status of hematopoietic stem cells as developmental intermediates between germ cells and MEL cells or by the presence of a second regulatory element, the viral long terminal repeat, within the oncoretrovirus vector construct. This element, which contains a potent enhancer and promoter that are transcriptionally active in primitive hematopoietic lineages, may serve to maintain the vector provirus in a relatively open chromatin conformation during the early stages of differentiation. This, in turn, would increase the chances that the accompanying HS-40/βpr/γ gene cassette would exist in a transcriptionally open chromatin state once the critical stage of erythropoiesis is reached. This latter interpretation is consistent with studies from others, indicating that linkage of recombinant β-globin and γ-globin genes to extended elements from the β-globin LCR with bone fide chromatin-opening activity are relatively resistant to gene silencing and are expressed at high levels following lentivirus-mediated transduction and transplantation of mouse bone marrow.[31–33] Future studies will be needed to determine whether addition of elements with chromatin-opening activity to our HS-40/βpr/γ gene cassette will lead to similar results in bone marrow transduction and transgenic mouse models.

It is also difficult to explain why the HS-40 element is able to activate expression of some genes but not others in transgenic mice. For example, minimal transgenes for α-globin,2 as well as LacZ or human growth hormone transcribed by a ζ-globin gene promoter,16,34 are efficiently activated when linked to HS-40 in transgenic mice. However, we failed to observe similar levels of activation when HS-40 was linked to the hybrid βpr/γ gene coding cassette. Other studies involving transduction of bone marrow with oncoretrovirus vectors suggest that HS-40 is also incapable of efficiently activating expression of a minimal coding cassette for γ-globin in vivo.[24] Taken together, these results suggest that there are inherent differences between the various transcriptional cassettes, such as the presence of specific CpG (cytosine-phosphate-guanosine) islands in the α-globin genes or as yet unidentified regulatory elements, which influence either the formation of silent heterochromatin or the ability to interact with the HS-40 element. Additional research will be required to distinguish between these 2 possibilities.

Finally, these studies provide an important insight into the strengths and weaknesses of the various methods used to study gene regulation. Many experimental systems are available for this purpose, including in vitro transcription, transient or stable transfection of somatic cells, and germ line transfer in mice. Likewise, many reporter gene systems have been designed to improve both the efficiency and sensitivity of detecting gene expression. In vitro transcription and transient transfection assays are popular because of their ease and rapidity. However, these systems generally rely on the use of naked DNA templates and as such are not suitable for studying gene activity regulated by chromatin. At the other extreme, transgenic mouse models provide the ability to analyze gene expression in a developmentally normal regulatory environment. However, such models are costly and time-consuming, making them impractical for detailed analysis of *cis* regulatory elements. As an intermediate approach, gene expression studies are also carried out by stable transfection of primary tissues or model cells lines. This approach allows for the relatively efficient and high-throughput analysis of engineered sequences in the context of intact chromatin and often serves as a good predictor of behavior in vivo. However, as demonstrated in the studies presented here, results observed by stable transfection in vitro do not always recapitulate results observed in vivo, indicating that care should be taken when interpreting such results. Indeed, differences in results observed in stably transfected cell lines and transgenic animals could provide important clues for studying how gene activity is regulated in vivo.

Acknowledgments

We thank Douglas R. Higgs for the α-globin HS-40 element and Gary Felsenfeld for the cHS4 chromatin insulator.

Supported by grants from the National Institutes of Health.

References

1. Higgs DR, Sharpe JA, Wood WG. Understanding alpha globin gene expression: a step towards effective gene therapy. Semin Hematol 1998;35:93–104. [PubMed: 9565152]
2. Sharpe JA, Chan-Thomas PS, Lida J, Ayyub H, Wood WG, Higgs DR. Analysis of the human alpha globin upstream regulatory element (HS-40) in transgenic mice. EMBO J 1992;11:4565–4572. [PubMed: 1425590]
3. Sharpe JA, Wells DJ, Whitelaw E, Vyas P, Higgs DR, Wood WG. Analysis of the human alpha-globin gene cluster in transgenic mice. Proc Natl Acad Sci U S A 1993;90:11262–11266. [PubMed: 8248238]
4. Higgs DR, Wood WG, Jarman AP, et al. A major positive regulatory region located far upstream of the human alpha-globin gene locus. Genes Dev 1990;4:1588–1601. [PubMed: 2253879]
5. Ren S, Wong BY, Li J, Luo XN, Wong PM, Atweh GF. Production of genetically stable high-titer retroviral vectors that carry a human gamma-globin gene under the control of the alpha-globin locus control region. Blood 1996;87:2518–2524. [PubMed: 8630419]
6. Li Q, Emery DW, Fernandez M, Han H, Stamatoyannopoulos G. Development of viral vectors for gene therapy of beta-chain hemoglobinopathies: optimization of a gamma-globin gene expression cassette. Blood 1999;93:2208–2216. [PubMed: 10090929]
7. Emery DW, Morrish F, Li Q, Stamatoyannopoulos G. Analysis of gamma-globin expression cassettes in retrovirus vectors. Hum Gene Ther 1999;10:877–888. [PubMed: 10223722]
8. Emery DW, Chen H, Li Q, Stamatoyannopoulos G. Development of a condensed locus control region cassette and testing in retrovirus vectors for A gamma-globin. Blood Cells Mol Dis 1998;24:322–339. [PubMed: 10087991]
9. Emery DW, Yannaki E, Tubb J, Nishino T, Li Q, Stamatoyannopoulos G. Development of virus vectors for gene therapy of beta chain hemoglobinopathies: flanking with a chromatin insulator reduces gamma-globin gene silencing in vivo. Blood 2002;100:2012–2019. [PubMed: 12200360]
10. Forrester WC, Novak U, Gelinas R, Groudine M. Molecular analysis of the human beta-globin locus activation region. Proc Natl Acad Sci U S A 1989;86:5439–5443. [PubMed: 2748594]

11. Li Q, Zhang M, Han H, Rohde A, Stamatoyannopoulos G. Evidence that DNase I hypersensitive site 5 of the human beta-globin locus control region functions as a chromosomal insulator in transgenic mice. Nucleic Acids Res 2002;30:2484–2491. [PubMed: 12034837]
12. Chung JH, Whiteley M, Felsenfeld G. A 5′ element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. Cell 1993;74:505–514. [PubMed: 8348617]
13. Li Q, Stamatoyannopoulos JA. Position independence and proper developmental control of gamma-globin gene expression require both a 5′ locus control region and a downstream sequence element. Mol Cell Biol 1994;14:6087–6096. [PubMed: 8065342]
14. Duan ZJ, Fang X, Rohde A, Han H, Stamatoyannopoulos G, Li Q. Developmental specificity of recruitment of TBP to the TATA box of the human gamma-globin gene. Proc Natl Acad Sci U S A 2002;99:5509–5514. [PubMed: 11960008]
15. Thorpe SJ, Thein SL, Sampietro M, Craig JE, Mahon B, Huehns ER. Immunochemical estimation of haemoglobin types in red blood cells by FACS analysis. Br J Haematol 1994;87:125–132. [PubMed: 7524614]
16. Huang BL, Fan-Chiang IR, Wen SC, et al. Derepression of human embryonic zeta-globin promoter by a locus-control region sequence. Proc Natl Acad Sci U S A 1998;95:14669–14674. [PubMed: 9843947]
17. Li Q, Peterson KR, Fang X, Stamatoyannopoulos G. Locus control regions. Blood 2002;100:3077–3086. [PubMed: 12384402]
18. West AG, Gaszner M, Felsenfeld G. Insulators: many functions, many mechanisms. Genes Dev 2002;16:271–288. [PubMed: 11825869]
19. Hatton CS, Wilkie AO, Drysdale HC, et al. Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human alpha globin gene cluster. Blood 1990;76:221–227. [PubMed: 2364173]
20. Jarman AP, Wood WG, Sharpe JA, Gourdon G, Ayyub H, Higgs DR. Characterization of the major regulatory element upstream of the human alpha-globin gene cluster. Mol Cell Biol 1991;11:4679–4689. [PubMed: 1875946]
21. Craddock CF, Vyas P, Sharpe JA, Ayyub H, Wood WG, Higgs DR. Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. EMBO J 1995;14:1718–1726. [PubMed: 7737123]
22. Barbour VM, Tufarelli C, Sharpe JA, et al. alpha-thalassemia resulting from a negative chromosomal position effect. Blood 2000;96:800–807. [PubMed: 10910890]
23. Emery DW, Yannaki E, Tubb J, Stamatoyannopoulos G. A chromatin insulator protects retrovirus vectors from chromosomal position effects. Proc Natl Acad Sci U S A 2000;97:9150–9155. [PubMed: 10908661]
24. Lung HY, Meeus IS, Weinberg RS, Atweh GF. In vivo silencing of the human gamma-globin gene in murine erythroid cells following retroviral transduction. Blood Cells Mol Dis 2000;26:613–619. [PubMed: 11358353]
25. Phillips RL, Ernst RE, Brunk B, et al. The genetic program of hematopoietic stem cells. Science 2000;288:1635–1640. [PubMed: 10834841]
26. Pannell D, Ellis J. Silencing of gene expression: implications for design of retrovirus vectors. Rev Med Virol 2001;11:205–217. [PubMed: 11479927]
27. Rivella S, Sadelain M. Genetic treatment of severe hemoglobinopathies: the combat against transgene variegation and transgene silencing. Semin Hematol 1998;35:112–125. [PubMed: 9565154]
28. Muller C, Leutz A. Chromatin remodeling in development and differentiation. Curr Opin Genet Dev 2001;11:167–174. [PubMed: 11250140]
29. Pannell D, Osborne CS, Yao S, et al. Retrovirus vector silencing is de novo methylase independent and marked by a repressive histone code. EMBO J 2000;19:5884–5894. [PubMed: 11060039]
30. Vassilopoulos G, Navas PA, Skarpidi E, et al. Correct function of the locus control region may require passage through a nonerythroid cellular environment. Blood 1999;93:703–712. [PubMed: 9885233]
31. May C, Rivella S, Callegari J, et al. Therapeutic haemoglobin synthesis in beta-thalassaemic mice expressing lentivirus-encoded human beta-globin. Nature 2000;406:82–86. [PubMed: 10894546]
32. Pawliuk R, Westerman KA, Fabry ME, et al. Correction of sickle cell disease in transgenic mouse models by gene therapy. Science 2001;294:2368–2371. [PubMed: 11743206]

33. Persons DA, Allay ER, Sawai N, et al. Successful treatment of murine beta-thalassemia using in vivo selection of genetically modified, drug-resistant hematopoietic stem cells. Blood 2003;102:506–513. [PubMed: 12663444]

34. Robertson G, Garrick D, Wu W, Kearns M, Martin D, Whitelaw E. Position-dependent variegation of globin transgene expression in mice. Proc Natl Acad Sci U S A 1995;92:5371–5375. [PubMed: 7777514]

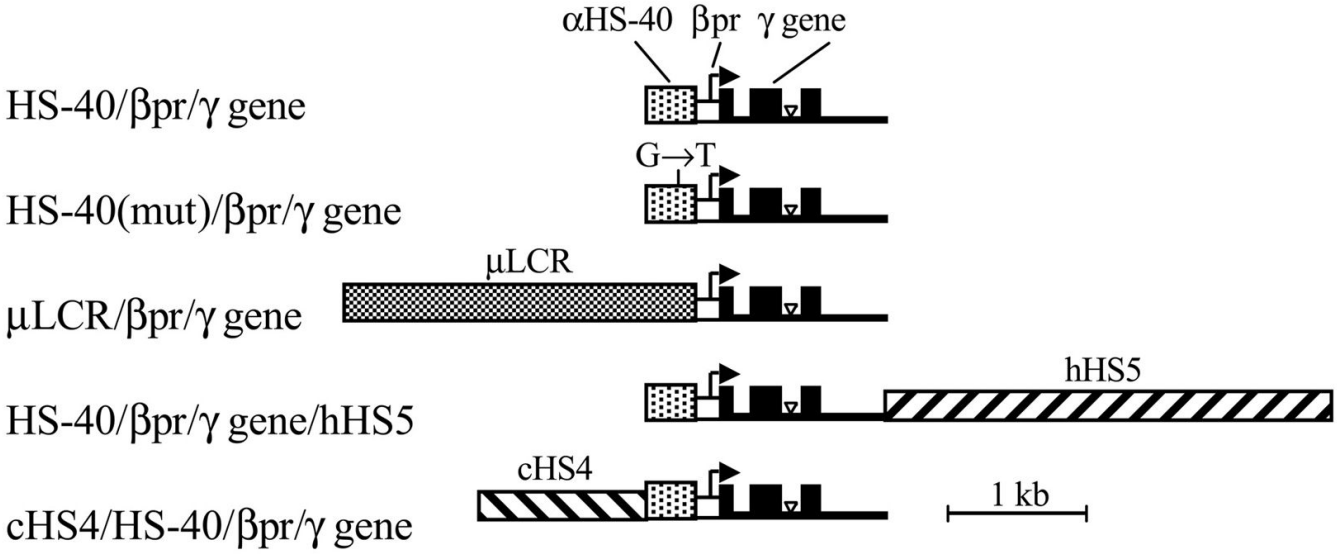

Figure 1. Transgene constructs

All transgene constructs contain a transcription unit that comprises a truncated β-globin promoter (βpr; □) and γ-globin coding segment (γ gene) containing a partial deletion (▽) of the second intron (■, exons; black line, untranslated regions). Construct HS-40/βpr/γ gene contains the HS-40 element from the human α-globin locus (αHS-40; ▩). Construct HS-40 (mut)/βpr/γ gene contains the HS-40 element with a G-to-T transversion in the 3′ NF-E2 (nuclear factor-erythroid 2) motif. Construct μLCR/βpr/γ gene contains a 2.5-kb composite μLCR (▩). Construct HS-40/βpr/γ gene/hHS5 contains HS-40 and a 3.2-kb fragment encompassing the hHS5 chromatin insulator (▨). Construct cHS4/HS-40/βpr/γ gene contains HS-40 and a 1.2-kb fragment encompassing the cHS4 chromatin insulator (▧).

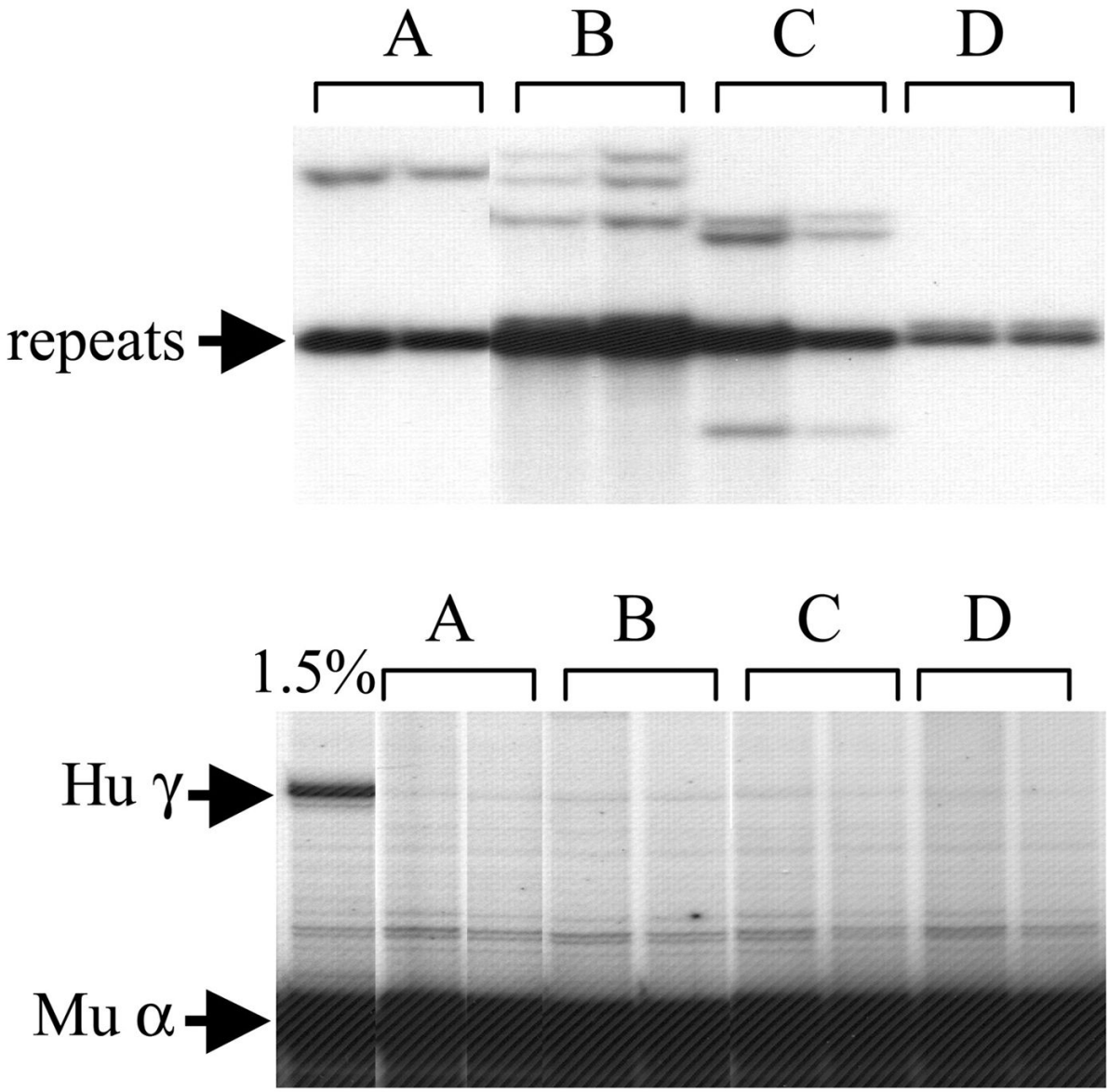

Figure 2. Analyses of structure and expression of construct HS-40/βpr/γ gene in transgenic mice

Four lines (A–D) carrying this transgene were established. (Top) Structure analysis. DNA from 2 individual animals of each line was digested with *Nco*I, which cuts once within the construct, and analyzed by Southern blotting with the use of a probe for HS-40. The arrow to the left indicates the 1.7-kb repetitive fragments generated from head-to-tail repeats of the construct. Other bands were from the end fragments containing the HS-40 fragment of the construct and various sequences from the mouse genome. (Bottom) RNA analysis. Total RNA was prepared from the blood of 3- to 4-week-old mice, and RNase protection assay was performed with the use of RNA probes for human γ-globin (Hu γ) and mouse α-globin mRNAs (Mu α). The leftmost lane contains control RNA from cells expressing human γ-globin mRNA at 1.5% of mouse 3-globin mRNA.

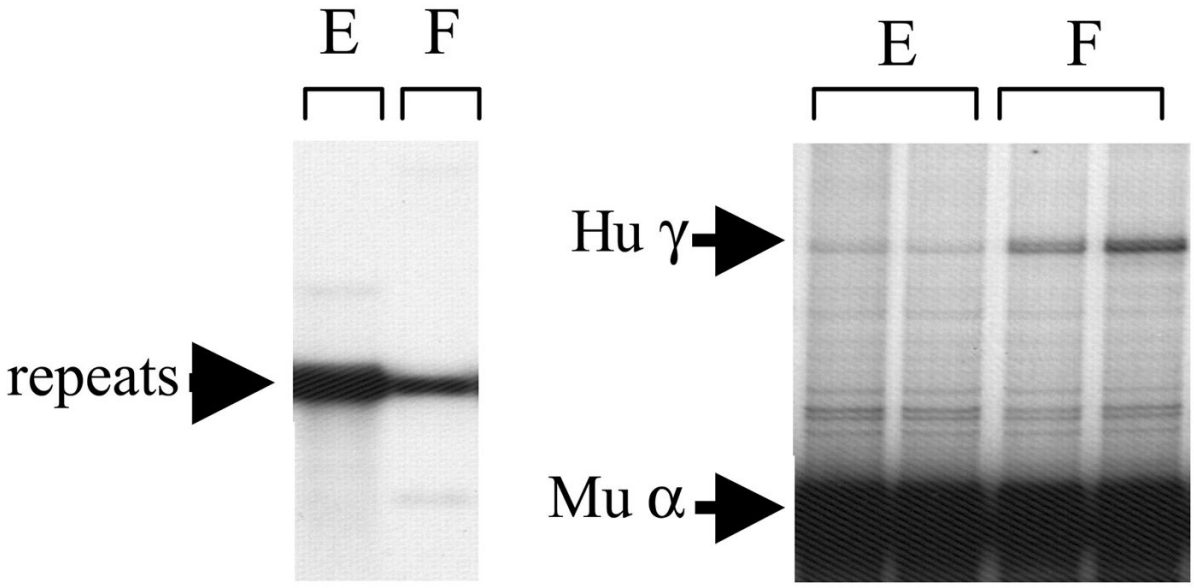

Figure 3. Analyses of structure and expression of construct HS-40(mutant)/βpr/γ gene in transgenic mice

Two lines (E–F) carrying this transgene were established. (Left) Structure analysis. (Right) RNA analysis. See the legend of Figure 2 for details.

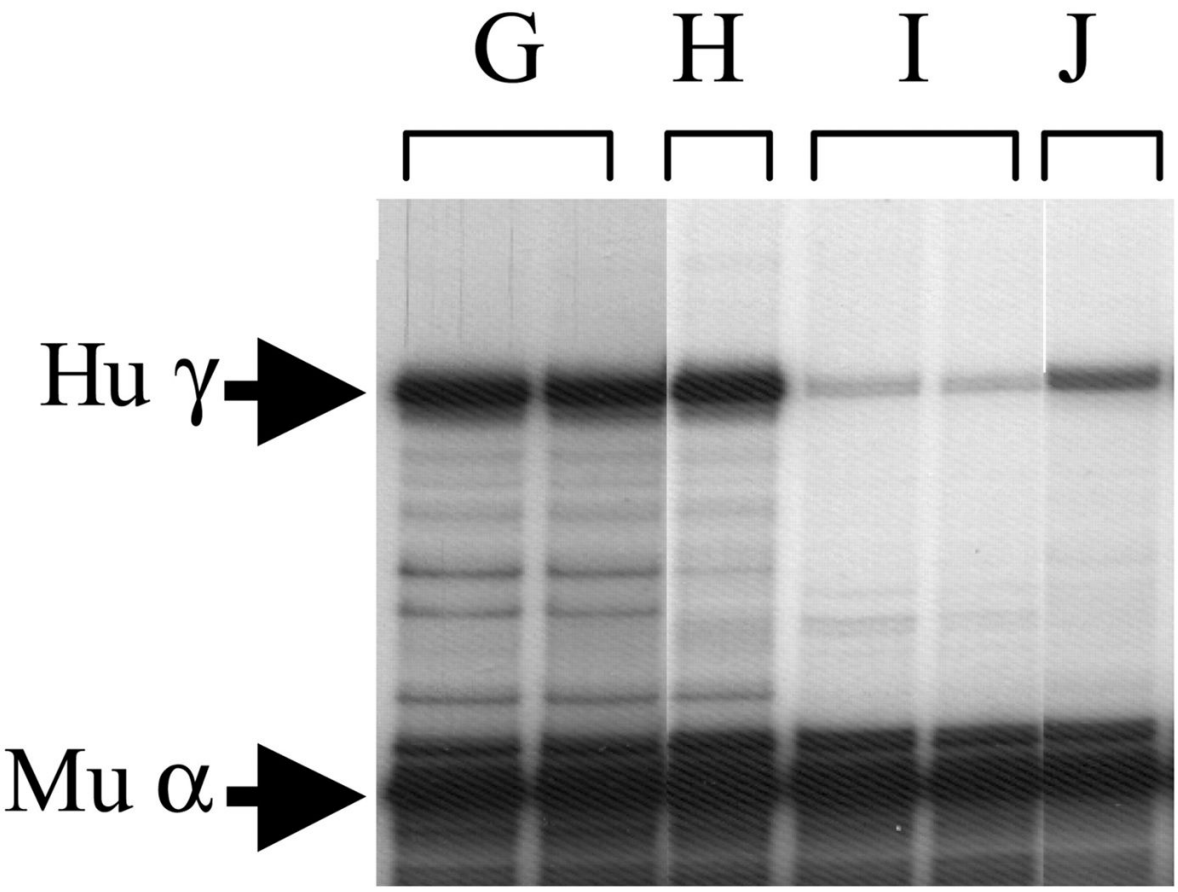

Figure 4. Analyses of structure and expression of construct μLCR/βpr/γ gene in transgenic mice
Four lines (G–J) carrying this transgene were established. RNA analysis is shown. See the legend of Figure 2 for details.

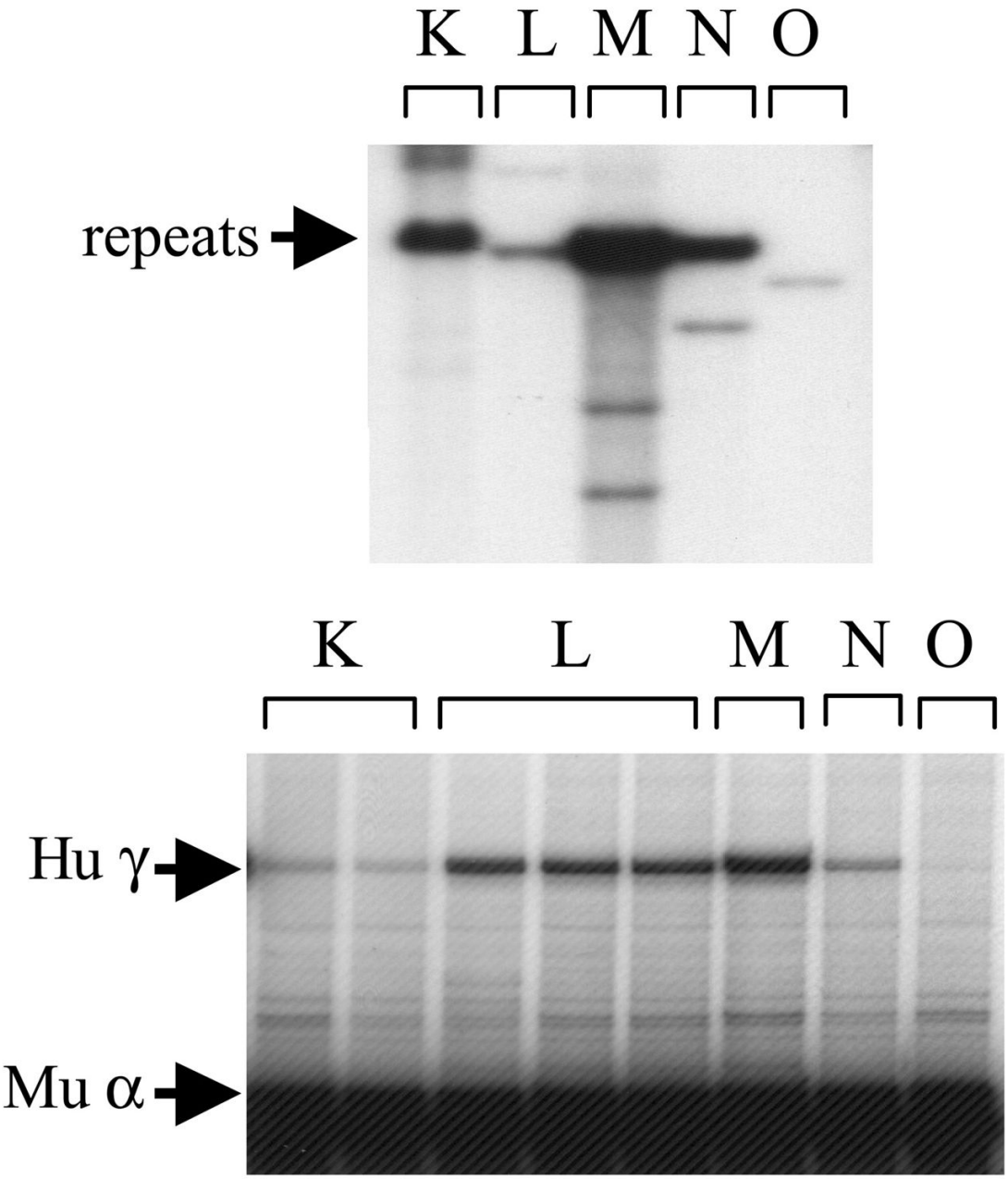

Figure 5. Analyses of structure and expression of construct HS-40/βpr/γ gene/HS5 in transgenic mice

Five lines (K–O) carrying this transgene were established. (Top) Structure analysis. (Bottom) RNA analysis. See the legend of Figure 2 for details.

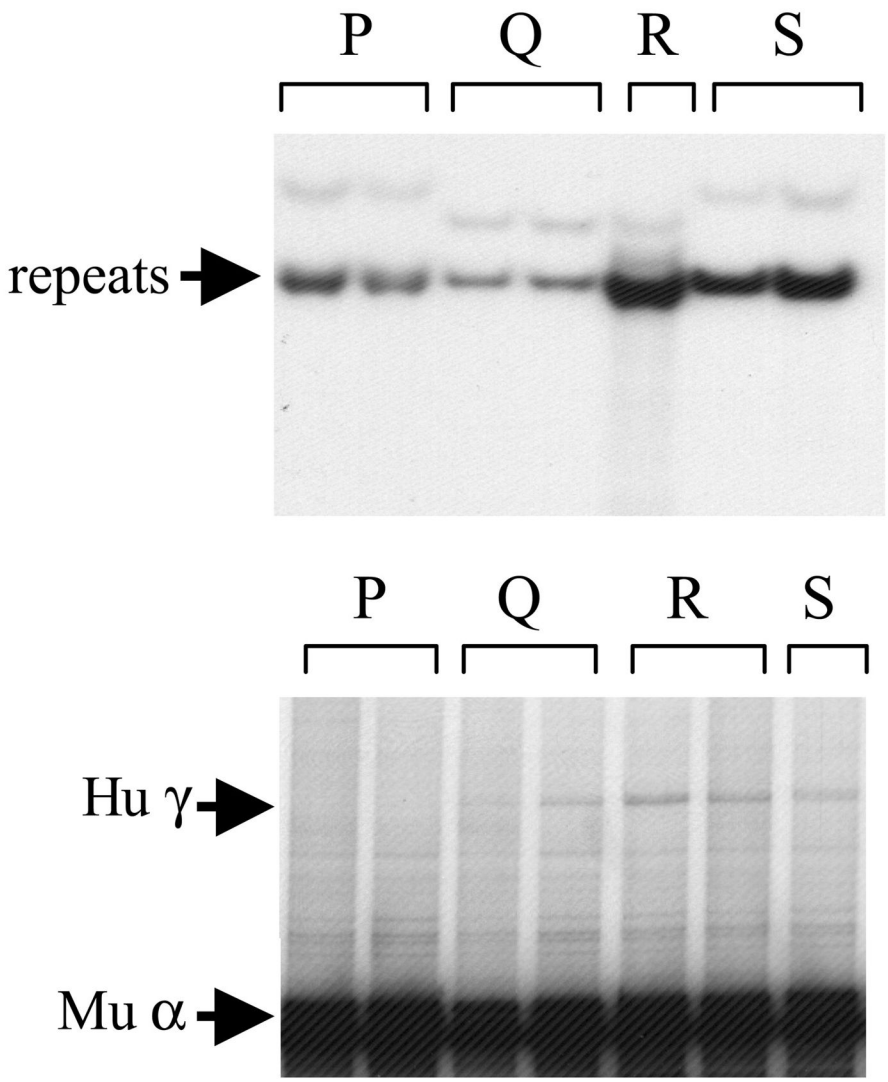

Figure 6. Analyses of structure and expression of the construct HS4/HS-40/βpr/γ gene in transgenic mice

Four lines (P–S) carrying this transgene were established. (Top) Structure analysis. (Bottom) RNA analysis. See the legend of Figure 2 for details.

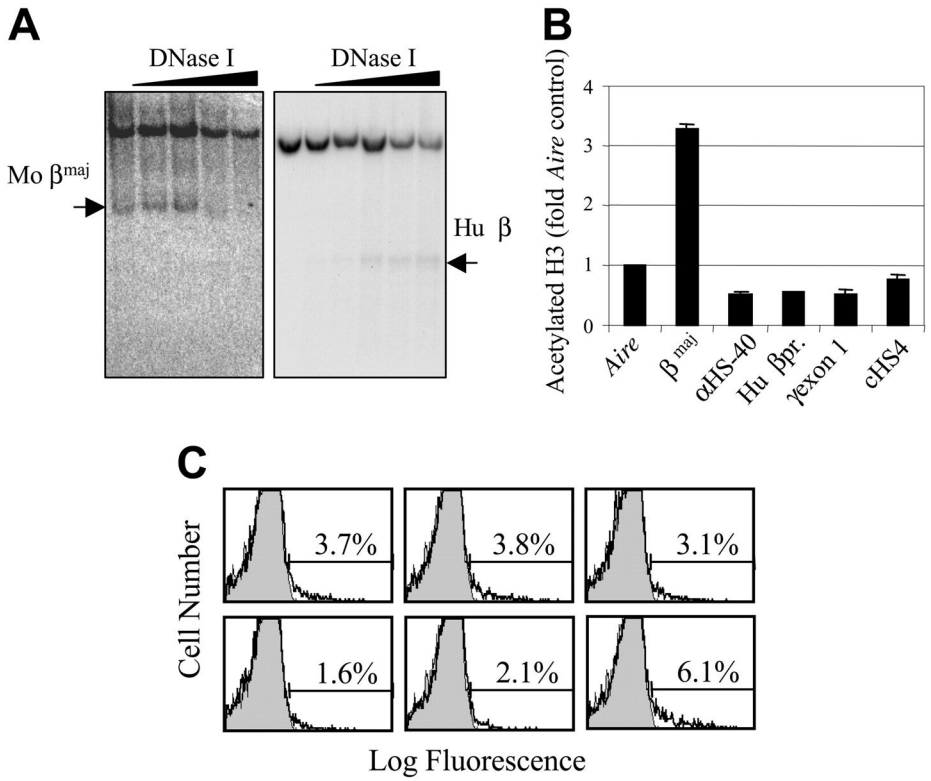

Figure 7. Nature of transgene silencing

(A) DNase I hypersensitivity assay. Nuclei from splenocytes containing the cHS4/HS-40/βpr/γ gene transgene were treated with increasing amounts of DNase I, and then purified DNA was digested completely with *Eco*RI (right) or *Sca*I (left), Southern blotted, and hybridized with a probe for 3′ γ-globin (right) or a probe for the mouse β^{maj} gene (left). The arrow indicates the subband corresponding to specific digestion at the human β gene promoter (Hu β) and the mouse β^{maj} gene (Mo β^{maj}). (B) Chromatin immunoprecipitation (ChIP) assay. Splenocytes containing the cHS4/HS-40/βpr/γ gene transgene were cross-linked with formaldehyde and sonicated, and the soluble chromatin was immunoprecipitated with an antibody against acetylated histone H3. The immunoprecipitated DNA was purified and quantified by real-time PCR with the use of primers specific for the indicated portions of the transgene, the endogenous mouse β^{maj} gene (as a positive control), and the endogenous mouse *Aire* gene (as an unexpressed negative control). All data are expressed as fold differences compared with the mouse *Aire* gene. Error bars indicate standard deviation (SD). (C) Immunofluorescent analysis of γ-globin transgene expression. Red blood cells (RBCs) from mice containing the HS4/HS-40/βpr/γ gene were fixed, permeabilized, stained with a phycoerythrin–conjugated anti-HbF antibody, and analyzed by flow cytometry. Profiles for transgenic animals (unfilled histogram with bold line) are overlaid on a profile from a nontransgenic control (filled histogram, thin line). The percentage of positive cells is presented above the indicated gates.

Table 1

Summary of structure and expression analyses of the transgenes in transgenic mice

Construct/Line	No. Repeats*	No. Ends*	No. copies	γRNA/mu αRNA**	γ RNA/mu αRNA/copy**
HS-40/βpr/γ gene					
A	4	1	5	0.07	0.06
B	22	4	26	0.13	0.02
C	12	3	15	0.08	0.02
D	2	1	3	0.05	0.06
HS-40(mut)/βpr/γ gene					
E	70	1	71	0.15	0.01
F	29	2	31	0.86	0.11
μLCR/βpr/γ gene					
G	—	—	3	34.2	45.6
H	—	—	12	30.5	10.2
I	—	—	3	1.6	2.1
J	—	—	2	7.6	15.3
HS-40/βpr/γ gene/hHS5					
K	51	1	52	0.3	0.02
L	2	1	3	2.7	3.6
M	49	2	51	3.4	0.42
N	8	1	9	0.6	0.27
O	0	1	1	0.1	0.38
cHS4/HS-40/βpr/γ gene					
P	9	1	10	0.03	0.01
Q	4	1	5	0.16	0.12
R	23	3	26	0.46	0.07
S	15	1	16	0.27	0.07

* —, not detected;

** mu, murine.