Persistent Fetal γ-Globin Expression in Adult Transgenic Mice following Deletion of Two Silencer Elements Located 3′ to the Human Aγ-Globin Gene

Maria Gazouli,[1,2] Eleni Katsantoni,[3] Theodoros Kosteas,[4] and Nicholas P Anagnou[1,2]

[1]Laboratory of Biology, University of Athens School of Medicine, Athens, Greece; [2]Laboratory of Cell and Gene Therapy, Centre of Basic Research II, Biomedical Research Foundation of the Academy of Athens, Athens, Greece; [3]Hematology Division, Biomedical Research Foundation of the Academy of Athens, Athens, Greece; and [4]Institute of Molecular Biology and Biotechnology, FORTH, Heraklion, Greece

Natural deletions of the human γ-globin gene cluster lead to specific syndromes characterized by increased production of fetal hemoglobin in adult life and provide a useful model to delineate novel *cis*-acting elements involved in the developmental control of hemoglobin switching. A hypothesis accounting for these phenotypic features assumes that silencers located within the Aγ- to δ-gene region are deleted in hereditary persistence of fetal hemoglobin (HPFH) and δβ-thalassemias, leading to failure of switching. In the present study, we sought to clarify the *in vivo* role of two elements, termed Enh and F, located 3′ to the Aγ-globin, in silencing the fetal genes. To this end, we generated three transgenic lines using cosmid constructs containing the full length of the globin locus control region (LCR) linked to the 3.3-kb Aγ-gene lacking both the Enh and F elements. The Enh/F deletion resulted in high levels of Aγ-globin gene expression in adult mice in all single copy lines, whereas, the LCR-Aγ single copy lines which retain the Enh and F elements exhibited complete normal switching of the fetal Aγ-gene. Our study documents directly for the first time the *in vivo* role of these two gene-proximal negative regulatory elements in silencing the fetal globin gene in the perinatal period, and thus these data may permit their eventual exploitation in therapeutic approaches for thalassemias.

INTRODUCTION

Human β-globin gene expression is regulated tightly during development and hematopoiesis. The human β-globin locus comprises five developmentally regulated genes (5′-ε-Gγ-Aγ-δ-β-3′) whose high level and stage-specific expression depends on interactions with the locus control region (LCR), consisting of five major DNaseI hypersensitive sites (Figure 1). The LCR activates β-globin gene transcription through direct interaction with promoter regions (1,2), and is a major determinant of the chromatin structure of the locus (3). Mice transgenic for the human β-globin locus express the human genes in a developmentally regulated manner (4,5). Specifically, the human globin genes undergo two developmental switches in their activation. Expression of the embryonic ε- and fetal γ-globin genes, first activated during primitive erythropoiesis in the embryonic yolk sac, switches to expression of the γ-genes at the start of definitive erythropoiesis in the fetal liver, with a small contribution by β-globin. The second switch occurs gradually around birth with the activation of the adult stage-specific δ- and β-globin genes, with δ-globin making a minor contribution, whereas γ-globin expression is gradually suppressed to very low levels (1%–2%) by the end of the first year of life.

The individual genes have been shown to be regulated by a complex interplay between *cis* regulatory elements, *trans*-acting factors, enhancer competition, and epigenetic mechanisms (6). Despite a remarkable amount of progress in this field, the exact mechanism(s) of globin gene silencing still is not understood completely.

Understanding the molecular basis of globin gene switching and discovery of strategies to efficiently express γ-globin genes in the adult is of particular interest, since reactivation of the fetal γ-globin genes in the adult has been shown to ameliorate the effects of hemoglobinopathies (7).

A series of well established quantitative trait loci accounting for 20%–50% of the fetal hemoglobin (HbF) variability in healthy adults, such as: a) the C→T single nucleotide polymorphism at position –158 of the Gγ-gene, creating a restriction site for the enzyme *Xmn*I (8);
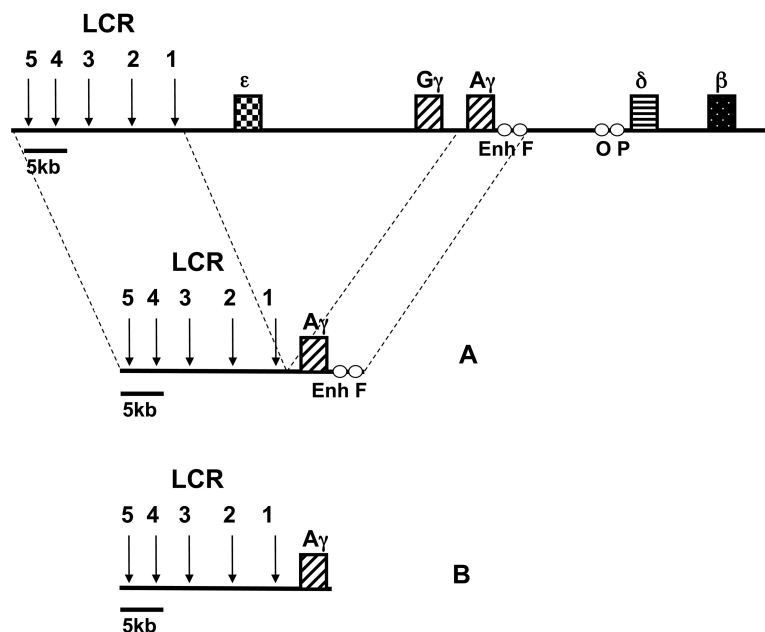
Figure 1. Schematic presentation of the constructs used in this study. (A) Control construct containing the LCR linked to the 5.6 kb Aγ-gene including both Enh and F elements. (B) ΔEnh/F construct containing the LCR linked to the 3.3 kb Aγ-gene lacking Enh and F elements. A diagram of the human β-globin locus is shown on the top of the panel, depicting the five functional globin genes and the four silencer elements Enh, F, O, and P. The vertical arrows denote the five major DNaseI hypersensitive sites of the LCR.

b) the HBS1L-MYB intergenic region on chromosome 6q23 (9); and c) the *BCL11A* gene on chromosome 2, encoding a zinc finger transcription factor acting directly within the β-cluster (10), has been studied extensively and its contribution on the mechanisms of fetal globin expression and hemoglobin switching has been evaluated recently (11). Furthermore, the configuration $(AT)_9T_5$ of the polymorphic sequence motif residing 0.5 kb 5' to the β-gene has been considered to confer high γ-globin levels in β-thalassemia patients in association with the *XmnI* polymorphism (12–14), although other studies (15) have found no effect. Furthermore, analysis of mutants with rare point mutations either in the Gγ or Aγ promoter associated with a hereditary persistence of fetal hemoglobin (HPFH) phenotype (11,16) have provided insights for the mechanisms of continued HbF synthesis based on the alterations in DNA-protein binding sites, resulting either in the creation of a new motif that allows

trans-activator binding, or in the abolishment of a repressor protein binding motif, leading to de-repression of fetal genes (11,16).

However, the most informative model for the role of *cis*-acting elements of the cluster involved in hemoglobin switching has been provided by a number of naturally occurring deletions in the locus leading to persistent γ-globin gene expression in the adult stage, and suggesting that γ-gene suppression is likely a complex process but one that can be perturbed (16). The persistence of fetal hemoglobin associated with such deletions is classified into two related clinical syndromes, that is, HPFH and (δβ)°-thalassemia (16). HPFH results in a substantial (14%–30%) pancellular γ-globin gene expression, whereas (δβ)°-thalassemias give rise to lower levels (2%–15%) of heterocellular γ-globin expression in the adult. In a particular set of deletions where the 3' end of the locus is lost, the 5' breakpoints map within a region between the Aγ- to

δ-globin genes. Based on the above observations, it has long been postulated that the Aγ- to δ-globin intergenic region harbors negative *cis*-acting elements involved in the regulation of the fetal-to-adult switch (17). It also has been suggested (18), without being mutually exclusive, that the juxtaposition of distal sequences located downstream of the 3' breakpoint of these deletions with enhancer-like function also may be involved in persistent γ-globin expression. The validity of the latter hypothesis for the HPFH deletions has been documented conclusively (experimentally) *in vivo* by us (19,20) and by others (21) with the identification and functional characterization of a series of HPFH enhancers (19–21). A similar mechanism may be operating in the HPFH-5 and HPFH Kenya deletions via the juxtaposition of the 3' β-globin enhancer to the proximity of the fetal genes (1,16).

Previous studies of our laboratory (22,23) have identified four elements, termed Enh, F, O, and P, located within the Aγ- to δ-globin intergenic region (see Figure 1) exhibiting silencer activity in transient transfection assays. Using electrophoretic mobility shift assays coupled with oligonucleotide competition and DNaseI footprinting techniques, several binding sites for the transcription factors YY1, GATA-1, and CP1 were identified in the F element. The six YY1 binding sites exhibit a variable degree of homology to the consensus motifs and bind YY1 with different affinity. Of the three strong YY1 sites, two sites residing at the 5' and 3' ends of the element, respectively, seem to confer its silencing activity (22,23). Based on these special features, these elements were considered as candidate regulatory *cis*-acting sequences for suppressing γ-globin expression in the adult stage.

Subsequent transgenic mouse studies using several of these elements have been controversial as to their *in vivo* function. Specifically, constructs that included the LCR linked to an Aγ-globin gene fragment containing Enh and F elements were silenced autonomously in

the adult stage of transgenic mice (24), indicative of their silencing activity. In contrast, deletion of a 12.5-kb region between the Aγ- to δ-globin genes, in a human β-globin locus yeast artificial chromosome (YAC) which includes all four elements, has been reported to have no effect on human globin gene switching in transgenic mice (25). Furthermore, deletion of Enh alone from a human β-globin locus YAC showed no observable effects on the regulation of the β-globin locus in transgenic mice (25,26).

Taken together, these findings, although not conclusive, do provide strong evidence that the identified elements possess transcriptional regulatory activity at specific developmental stages and may be involved in the developmental regulation of globin gene switching. Supporting this hypothesis, a recent study from our group (27) documented that deletion of both Enh and F elements in the context of a 185-kb human β-globin locus PAC (P1 phage artificial chromosome) results in an increase of transcription of both ε- and γ-globin genes in the embryonic yolk sac stage, whereas no effects on human globin gene expression in the fetal liver and adult blood stages were observed. The latter finding apparently reflects a known phenomenon of functional redundancy (28), since the rest of the so far identified silencer elements (that is, O and P elements and possibly others) remained intact and thus functional in our modified PAC, actually masking the net effect of the absence of Enh and F elements on the fetal γ genes (27).

In view of these unsettled data, in the present study, we sought to further clarify the *in vivo* function of the Enh and F elements in the context of cosmid constructs containing the human globin LCR directly linked to the Aγ gene, either without or in the presence of the two silencers retained in their physical position. Our goal was to delineate the direct effect of these elements on the Aγ-gene silencing during the perinatal period, in the absence of any other silencer elements of the human β-globin locus.

MATERIALS AND METHODS

DNA Constructs

The cosmid construct used was generated by ligation and packaging into phage of the following fragments: a) a 27-kb *Pvu*I-*Cla*I fragment from LCR-loxP-Aγ-3′HS cosmid (20) containing the cos site of the pTCF cosmid vector (1), the LCR, and the loxP site; b) a *Cla*I-*Kpn*I fragment from a pBluescript-Aγ plasmid containing the Aγ gene (3.3 kb); c) a 14-kb *Kpn*I-*Not*I fragment from the LCRε cosmid (4) containing a cos site and sequences downstream of the ε-globin gene. Packaging was performed using Gigapack Gold extracts (Stratagene, La Jolla, CA, USA) according to the manufacturer's instructions.

Transgenic Mice

The fragment containing the 25-kb LCR-loxP-Aγ insert (see Figure 1) was released by *Sal*I–*Kpn*I digestion, purified by Elutip D (Schleicher & Schuell, BioScience, Bath, UK), checked for DNA integrity and concentration, and prepared for microinjection, essentially as described previously (27). Purified DNA was injected into the pronucleus of fertilized eggs of FVB/N mice as described previously (27). Transgenic founders were identified by polymerase chain reaction (PCR) using the primers in the HS5 site of the LCR and by Southern blotting using the LCR HS5 3.3-kb *Eco*RI fragment and the 5′ Aγ 1.7-kb *Eco*RI-*Bam*HI fragment as probes. The integrity of cosmid transgenes was checked by cosmid hybridization according to Strouboulis *et al*. (4). Transgene copy numbers were determined by using a 5′ Aγ 1.7-kb *Eco*RI-*Bam*HI fragment and a 0.9-kb *Pvu*I fragment from the endogenous mouse carbonic anhydrase II (*CA2*) gene, and the ratios of the intensities of the two probes were compared with those obtained for the single copy LCR-Aγ-HPFH2C transgenic line (20). PhosphorImager analysis was performed using ImageQuant software (Molecular Dynamics, Sunnyvale, CA, USA). Multicopy transgenic mouse lines were bred

with the CAG-Cre transgenic lines (29) to generate single copy animals. The integrity and the copy number of the single copy mice were determined similarly as for the multicopy lines. The control transgenic mice used in this study were the LCR-loxP-Aγ lines generated by our group (20). All experiments were performed according to the guidelines of the Animal Facility of IMBB, as approved by the Institutional Committee.

S1 Nuclease Protection Assays

S1 nuclease protection analysis was carried out using total RNA from blood of adult animals. RNA was isolated using the Trizol reagent according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). The probes used and conditions for S1 nuclease protection assays and polyacrylamide gel electrophoresis were essentially as described previously (30–32). Specific activities of probes were determined as described previously (32), and are indicated in the legend of Figure 2. Quantitation of expression levels was performed on a PhosphorImager using the ImageQuant software (Molecular Dynamics, Sunnyvale, CA, USA).

Real-Time Reverse Transcriptase PCR

Total RNA was isolated from adult blood using Trizol reagent (Invitrogen) and treated with RQ1 RNase-Free DNase (Promega, Madison, WI, USA) for 30 min at 37°C according to the manufacturer's instructions. DNase-treated RNA samples were reverse transcribed with Superscript II Reverse Transcriptase (Invitrogen) using Oligo dT$_{16}$. Control reactions without reverse transcriptase also were performed. For the real-time PCR assay, primer sets specific for the human γ-globin gene (forward primer exon 2, reverse primer exon 3) were used. Primer sequences are available upon request. Normalization for the amount of template was performed using primers specific for exon 8 of the mouse *HPRT* gene. Amplified products were approximately 100 nt long. Reactions were performed in duplicate, using the SYBR Green PCR
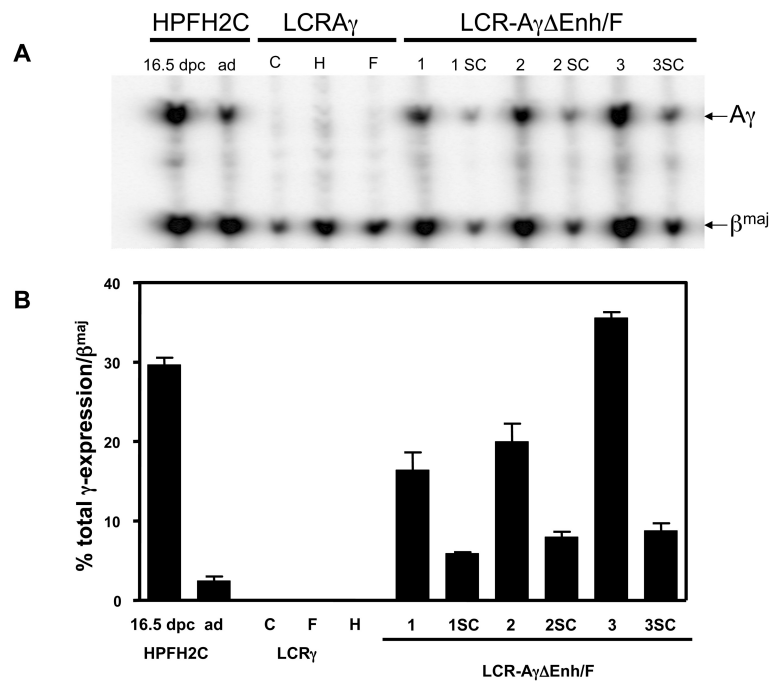
Figure 2. Expression of Aγ-globin gene in multicopy and single-copy ΔEnh/F lines. (A) Representative S1 nuclease protection analysis of expression of human Aγ versus mouse β^maj globin genes in adult blood RNA of single copy control lines, multicopy LCR-Aγ-ΔEnh/F-1, 2, and 3 lines, and the derived single copy (sc) lines analyzed. The type of protected fragments are indicated to the right of the panel. Controls used in this experiment included a 16.5 dpc embryo from HPFH2C line and an adult HPFH2C RNA sample (ad), expressing high levels of the Aγ-gene, as described by Katsantoni *et al.* (20). Expression levels were calculated by ImageQuant after correcting for probe-specific activities according to the ratio of 1:2.4 for Aγ:β^maj probes. (B) Quantitation of Aγ-globin expression levels of the multicopy LCR-Aγ-ΔEnh/F-1, 2, and 3 lines, and the derived single copy (sc) lines as a percentage of total β^maj globin gene expression. Data shown represent mean ± standard deviation from a total of three samples analyzed per each line. The values used in the chart are from Table 1.

Master Mix (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions. Real-time PCR was performed in an ABI PRISM 7000 Sequence Detection System (Applied Biosystems), as follows: initial denaturation for 2 min at 50°C and for 10 min at 95°C, followed by 40 cycles of PCR (95°C for 15 sec; 60°C for 1 min). Data were analyzed using the comparative C_T method for the relative quantitation of results (33). Post-amplification denaturation curves showed that the primer pairs generated single products.

DNA Fluorescence *in situ* Hybridization (FISH) Analysis

Peripheral blood cells were cultured for 72 h in RPMI 1640 medium (Invitrogen,

Carlsbad, CA, USA). Chromosome preparations were made according to standard procedures. Fluorescence *in situ* hybridization (FISH) was carried out as described by Mulder *et al.* (34). The probe used to detect the transgene was the biotin-labeled 5' Aγ 1.7-kb *Eco*RI-*Bam*HI fragment, which was detected immunochemically with fluorescein. Chromosomal DNA was counterstained with DAPI (4′,6 diamidino-2-phenylindole), which stains centromeric domains more intensely.

Linear Amplification Mediated-Polymerase Chain Reaction (LAM-PCR)

LAM-PCR was performed using the previously reported methodology (35). This newly established approach permits

the localization of the integration site of a construct in a given genome by identifying the sequence of an unknown genomic region flanking a known genomic segment. It represents a combination of linear amplification of target DNA with solid-phase second strand synthesis, followed by ligation of an oligonucleotide cassette, and then by nested exponential PCR. Briefly, genomic DNA from mouse tails was used as a template. Linear amplification was carried out with the biotinylated primer KN872_B (5′-TGAAG ACCTGGGGGCTGGATT-3′). The single stranded run-off DNA fragments were purified by magnetic capture. The single stranded DNA was made double stranded with random hexamers (Roche, Indianapolis, IN, USA) and the DNA polymerase Klenow fragment. The products were digested by *Tsp501I* and then ligated into a linker cassette which was generated by annealing two oligonucleotides: 5′-GACCCGGGAGATCTG AATTCAGTGGCACAGCAGTTAGG-3′ and 5′-AATTCCTAACTGCTTGCCAC TGAATTCAGATCTCCCGGGTC-3′. The resultant ligated products were exponentially amplified twice using the following primers: KN87-N1: 5′-ACCTGGGGGC TGGATTGATTG-3′ and 5′-GGGGCTGGAT TGATTGCAGCT-3′. The cycle sequencing of the specific LAM-PCR amplicon was performed using an ABI Prism genetic analyzer 3100 (Applied Biosystems), according to the manufacturer's instructions. The obtained sequence data were analyzed and mapped to the appropriate mouse chromosome locations, employing the Basic Local Alignment Search Tool (BLAST), developed by the National Center for Biotechnology Information (NCBI).

RESULTS

Generation of LCR-Aγ-ΔEnh/F Transgenic Lines

To investigate the effects of the deletion of Enh and F elements on the silencing of the Aγ-globin gene, we generated transgenic lines by microinjecting a 25-kb construct containing the full

length LCR and the 3.3-kb Aγ-globin gene lacking the Enh and F elements (LCR-Aγ-ΔEnh/F) as shown in Figure 1. Four founders transgenic for the LCR-Aγ-ΔEnh/F construct were obtained. The first founder mouse never passed on the transgene. The other three LCR-Aγ-ΔEnh/F founder mice successfully transmitted the transgene to establish the corresponding transgenic lines LCR-Aγ-ΔEnh/F-1, LCR-Aγ-ΔEnh/F-2, and LCR-Aγ-ΔEnh/F-3. Further analysis showed that all three lines contained an intact transgene (data not shown) in multiple copies: LCR-Aγ-ΔEnh/F-1, two copies; LCR-Aγ-ΔEnh/F-2, three copies; and LCR-Aγ-ΔEnh/F-3, four copies (as shown in Table 1).

Aγ-Gene Expression Analysis in Adult LCR-Aγ-ΔEnh/F Multicopy Lines

We first analyzed Aγ-globin gene expression levels against those of the endogenous β^{maj}-globin gene in RNAs isolated from adult blood of transgenic LCR-Aγ-ΔEnh/F multicopy lines by S1 nuclease protection assay (see Figure 2). As negative controls for this analysis, three LCR-Aγ single copy lines were used (lines C, F, and H), which retain the Enh and F elements downstream of the Aγ gene and exhibit normal switching, as reported previously by our group (20). Expression of Aγ-globin normally is switched off at this stage in these mice (20,24,28). This was corroborated in the three LCR-Aγ control lines (Table 1; see Figure 2). In contrast, all three multicopy LCR-Aγ-ΔEnh/F lines lacking the Enh and F silencers exhibited persistent Aγ-globin expression in the adult blood (see Figure 2) ranging from 13.31 ± 2.63 to 17.78 ± 0.62 per transgene copy and per copy of endogenous mouse β^{maj}-globin (Table 1). RNA samples from three different adult animals from each of the three generated LCR-Aγ-ΔEnh/F lines were analyzed. The highest LCR-Aγ-ΔEnh/F Aγ-expressing line was the LCR-Aγ-ΔEnh/F-3 (Table 1; see Figure 2).

We further analyzed Aγ-gene expression by quantitative real-time reverse transcriptase PCR. These results con-

firmed the S1 nuclease mapping data, with high γ-globin expression being detected in the adult blood of the LCR-Aγ-ΔEnh/F transgenic mice only, ranging from 24 ± 3 to 112 ± 4% of the levels of the LCR-Aγ-HPFH2C transgene at the fetal stage (20), which was used as a positive control (Table 2). On the contrary, no expression was detected in the adult blood of the control LCR-Aγ lines C and H containing the two silencer elements (Table 2). Since the only difference between the LCR-Aγ-ΔEnh/F *versus* the LCR-Aγ constructs is the selective absence or presence of the silencers, respectively, these findings directly demonstrate that the absence of Enh and F elements in the context of the LCR-containing transgene can alter the developmental expression of the Aγ-gene in

adult blood and leads to its persistent and efficient expression.

Persistent Adult Stage Aγ-Globin Expression in Single Copy LCR-Aγ-ΔEnh/F Lines

To distinguish whether the persistence of Aγ-gene expression in the LCR-Aγ-ΔEnh/F lines was due specifically to the absence of the Enh/F silencers or to a non-specific effect of multicopy integrants of LCR-containing globin transgenes, we generated single copy mice from all three multicopy animals by crossbreeding with lines expressing the Cre recombinase before the two-cell stage of embryonic development using the CAG-Cre mice (29). Three single copy (sc) LCR-Aγ-ΔEnh/F lines were generated and were designated as LCR-

Table 1. Summary of the control and multicopy lines generated, transgene copy numbers, γ-globin mRNA expression levels in adult blood, and integration sites of the transgene.

Transgenic lines	Copy no.	% γ/β^{maj} Globin[a]	% γ/β^{maj} Globin per copy[a]	Chromosomal integration site
HPFH2C 16.5 dpc embryo	1	26.33 ± 4.73	52.67 ± 9.45	Euchromatic
HPFH2C adult	1	2.46 ± 0.93	4.91 ± 1.87	Euchromatic
LCRAγ-C	1	0	0	Euchromatic
LCRAγ-F	1	0	0	Centromeric
LCRAγ-H	1	0	0	Euchromatic
LCR-Aγ-ΔEnh/F-1	2	16.40 ± 3.86	16.40 ± 3.86	Centromeric
LCR-Aγ-ΔEnh/F-2	3	19.97 ± 3.95	13.31 ± 2.63	Euchromatic
LCR-Aγ-ΔEnh/F-3	4	35.57 ± 1.25	17.78 ± 0.62	Euchromatic

[a]Expression of the fetal Aγ-gene was calculated by S1 nuclease analysis as 1) a percentage of total β^{maj} globin expression and 2) per transgene copy and per copy of the endogenous β^{maj} mouse globin gene. Mean and standard deviations are derived from three independent RNA samples.

Table 2. Real-time PCR analysis of γ-globin expression in adult blood of the generated multicopy transgenic lines.

Transgenic lines	% Expression γ-globin/γ-globin of 16.5 dpc embryo normalized to HPRT[a]
HPFH2C 16.5 dpc embryo	100
LCRAγ-C	0
LCRAγ-H	0
LCR-Aγ-ΔEnh/F-1	24 ± 3
LCR-Aγ-ΔEnh/F-2	66 ± 16
LCR-Aγ-ΔEnh/F-3	112 ± 4

[a]Normalization for the amount of template was performed using primers specific for exon 8 of the mouse *HPRT* gene. The results shown (mean ± standard deviation) represent the average of at least three independent RNA samples per each line.

Aγ-ΔEnh/F-1sc, LCR-Aγ-ΔEnh/F-2sc, and LCR-Aγ-ΔEnh/F-3sc. We then analyzed Aγ-gene expression in the adult blood by S1 nuclease protection. As controls for normal switching, we included the three single copy LCR-Aγ lines mentioned above. As expected, Aγ-gene expression was switched off in all three LCR-Aγ single copy control lines analyzed. In contrast, Aγ-globin was expressed in the adult blood of all three LCR-Aγ-ΔEnh/F single copy lines (Table 3). The levels of expression were slightly lower and copy number-dependent, compared with the corresponding multicopy lines, as expected (Table 3; see Figure 2). These results were confirmed independently by quantitative real-time reverse transcriptase PCR analysis (Table 4).

Correlation of Persistent Aγ-globin Expression with Integration Sites

Our data demonstrate that the deletion of Enh and F elements can lead to persistent Aγ-expression in adult mice carrying either single or multiple copies. We wished to further investigate whether this persistent Aγ-globin gene expression in the adult stage is linked to or is affected by specific chromosomal integrations. Position effects dependent on the chromatin structure at the sites of transgene integration often are observed (30). Therefore, we initially mapped transgene integration sites in all transgenic lines using FISH analysis in metaphase spreads to exclude the possibility that Aγ-gene expression could reflect position effects and not the actual effect of the absence of Enh and F silencer elements. We classified chromosomal sites as centromeric and euchromatic, the latter including all integrations that did not map close to a telomere or a centromere. All integration sites are shown in Figure 3 and the results are summarized in Table 1. The data documented that LCR-Aγ-ΔEnh/F-1 line is integrated in a centromeric region, whereas lines 2 and 3 contain euchromatic integration sites (Table 1; see Figure 3).

Table 3. Summary of the single copy (sc) lines generated, transgene copy numbers, γ-globin mRNA expression levels in adult blood, and integration sites of the transgene.

Transgenic lines	Copy no.	% γ/β^{maj} Globin^a	% γ/β^{maj} Globin per copy^a	Chromosomal integration site
LCR-Aγ-ΔEnh/F-1sc	1	5.90 ± 0.26	11.80 ± 0.52	Centromeric
LCR-Aγ-ΔEnh/F-2sc	1	7.97 ± 1.15	15.94 ± 2.30	Euchromatic
LCR-Aγ-ΔEnh/F-3sc	1	8.75 ± 1.64	17.50 ± 3.28	Euchromatic

^{a}Expression of the fetal Aγ-gene was calculated by S1 nuclease analysis as 1) a percentage of total β^{maj} globin expression and 2) per transgene copy and per copy of the endogenous β^{maj} mouse globin gene. Mean and standard deviations are derived from three independent RNA samples.

To further determine the exact genomic locations of the integration sites, we applied the recently introduced LAM-PCR method (35), which confirmed the FISH results (Table 5). The transgene of the LCR-Aγ-ΔEnh/F-1 line was integrated into chromosome 8 within the *neuropilin-1* and *integrin β1* gene locus (chr8qE2), while the transgenes of the LCR-Aγ-ΔEnh/F lines 2 and 3 were integrated near the *KIF27* gene (chr13qB2) and the *tyrosine phosphatase receptor type L* gene (chr4qD2.3), respectively.

In the case of the integration sites near the *integrin β1*, *KIF27*, and *tyrosine phosphatase receptor type L* genes, their transcription sites were located far enough from the integrated transgene (>100 kb of surrounding region), suggesting that there should be little if any effect from their transcriptional activation. However, we cannot exclude the theoretical possibility that the expression levels in LCR-Aγ-ΔEnh/F-1 line could be influenced by the *neuropilin-1* gene, since its transcrip-

tion site was found approximately 400 bp upstream of the Aγ-gene, while it has been shown previously to be expressed in CD45$^+$ hematopoietic cells from murine fetal liver (36).

The centromeric integration site in line LCR-Aγ-ΔEnh/F-1 is correlated with slightly lower Aγ-gene expression levels compared with lines 2 and 3 associated with euchromatic integration sites. The fact that Aγ-gene is expressed even when the transgene was integrated into the centromere, excludes the possibility that the Aγ-gene expression is due to positive position effects usually associated with open chromatin euchromatic sites, and strengthens the possibility that the resulting persistent Aγ-gene expression reflects actually altered developmental regulation of the gene, due solely to the absence of the Enh and F silencers.

DISCUSSION

Several studies have shown that, in murine models, individual human γ- and

Table 4. Real-time PCR analysis of γ-globin expression in adult blood of the generated single copy (sc) transgenic lines.

Transgenic lines	% Expression γ-globin/γ-globin of 16.5 dpc embryo normalized to HPRT^a
HPFH2C 16.5 dpc embryo	100
LCRAγ-C	0
LCRAγ-H	0
LCR-Aγ-ΔEnh/F-1sc	18 ± 1
LCR-Aγ-ΔEnh/F-2sc	45 ± 2
LCR-Aγ-ΔEnh/F-3sc	43 ± 5

^{a}Normalization for the amount of template was performed using primers specific for exon 8 of the mouse *HPRT* gene. The results shown (mean ± standard deviation) represent the average of at least three independent RNA samples per each line.
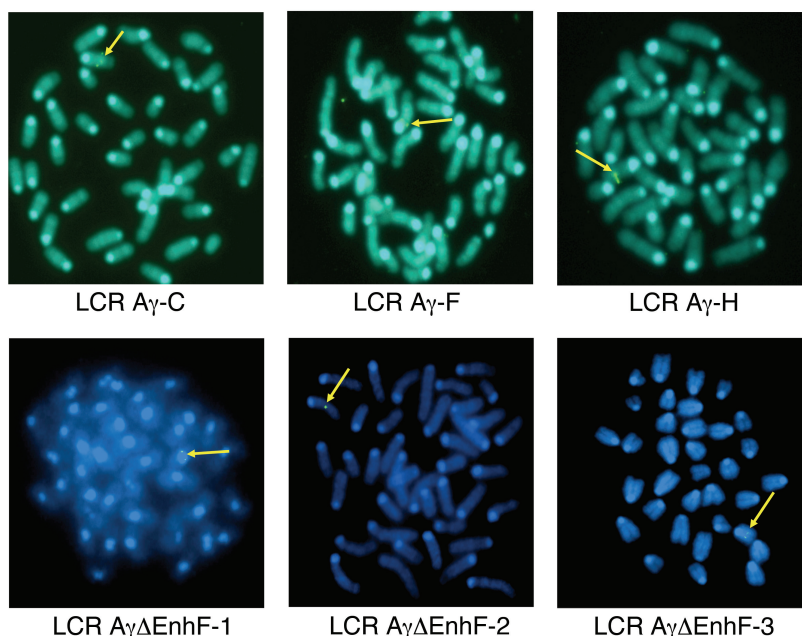
Figure 3. Chromosomal sites of transgene integration analyzed by FISH. Sites of chromosomal transgene integration are shown (arrow) for the control (LCR Aγ-C, LCR Aγ-F, and LCR Aγ-H) and for the single copy (LCR-Aγ-ΔEnh/F-1, 2, and 3) lines used in this study.

β-globin transgenes are expressed at high levels in the presence of the LCR, and are regulated appropriately at the fetal and adult stages of development (37,38). The δβ and Aγδβ-thalassemias and the deletion forms of hereditary persistence of fetal hemoglobin (HPFH) are naturally occurring mutations associated with persistent expression of fetal hemoglobin in adult life, albeit at variable levels. It has been proposed that the Aγ-δ globin intergenic sequences harbor negative regulatory elements that are involved in suppressing γ-globin expression in the adult stage (17). In accordance to this notion, we previously described two elements Enh and F, located 3' to the Aγ gene as silencers in transient transfection assays (22,23).

To further elucidate the direct *in vivo* role of the Enh and F elements on Aγ-globin

gene silencing under the control of LCR, we generated transgenic mice using cosmid constructs containing the full length LCR linked to the 3.3-kb Aγ-gene, but lacking selectively both the Enh and F silencers. The deletion of Enh and F elements resulted in persistent high levels of Aγ-expression in the adult stage in all three multicopy and three single copy lines tested. Furthermore, according to our results of FISH analysis, the deletion of Enh and F elements can lead to persistent Aγ-gene expression in the adult stage in all types of chromatin environments, not only in "permissive" ones.

Our observations are in agreement with earlier results using μLCR 3.3-kb Aγ (39) and LCR HS2 3.3-kb Aγ lines (40) exhibiting adult stage Aγ-gene expression in constructs lacking the region harboring the Enh and F silencers, and in contrast to earlier reports describing Enh initially as an enhancer element in transient reporter assays using a different setting (41). Additional studies by the same group have documented the mixed functional features of the Enh element (42).

Our previous experiments (27) have documented that the selective deletion of Enh and F elements from a 185-kb human β-globin locus PAC results in an increase of ε and γ-globin mRNA levels in the embryonic stage (that is, yolk sac stage) of erythropoiesis. This phenomenon was shown to be due to an increase

Table 5. The precise transgene integration sites identified by LAM-PCR in LCR-Aγ-ΔEnh/F lines used in this study.

Transgenic line	Sequence (5' → 3')[a]	Locus	Position to TSS (bp)[b]	Gene[c]	E-value[d]
LCR-Aγ-ΔEnh/F-1	GACCTGTAGCCCAGTGCCCAGAGCATATTATCATAACCAC ATTTCAGGGGACGCCAACGT	chr8qE2	>100 >100,000	5' side: *neuropilin-1* 3' side: *integrin β1*	9e–23
LCR-Aγ-ΔEnh/F-2	GAATTCCAGCACACTGGCGGCCGTTACTAGTGGATCCGAG	chr13qB2	>100,000	Kinesin-related protein 27 (*KIF27*)	1e–10
LCR-Aγ-ΔEnh/F-3	GACTCAGTTTCTTCATCTGTAGAATGG	chr4qD2.3	>100,000 >100,000	5' side: *hypothetical protein* 3' side: *tyrosine phosphatase, receptor type L*	2e–04

[a]The sequences shown are part of the LAM-PCR amplicons located 5' of the linker cassette.
[b]TSS, transcription start site.
[c]Precise chromosomal locations were defined by BLAST search for the mouse genome.
[d]E-value, expectation value (by BLAST search).

in the rate of transcription rather than to an increase in the number of cells transcribing the human globin locus (27). However, the human developmental switching from fetal γ-globin to adult β-globin gene expression was not affected by this particular deletion, thus identifying Enh and F as locus-wide regulatory elements capable of downregulating transcription of the human β-globin locus in an embryonic-specific manner. These results however, also implied the operation of a well known mechanism of functional redundancy (28), and, in the context of the full human β-globin locus, additional regulatory elements in the region between Aγ and δ genes (that is, O and P silencer elements we have described previously [22,23] located upstream of the δ gene) or in other regions of the cluster, might be needed to act synergistically with Enh and F elements to silence γ-globin genes in the adult stage of transgenic mice. This notion is corroborated by the phenotype of Corfu (δβ)° thalassemia, a natural deletion of a similar size and location, removing O and P elements, leading to significant levels of γ-globin in heterozygotes (43).

Nevertheless, in our previous study (20), we showed that in the presence of LCR and the two Enh and F silencers, and in the absence of additional regulatory elements or competition with the β-globin gene, Aγ-globin gene is silenced autonomously in the adult.

Furthermore, elements of the proximal and the distal Aγ-gene promoter (44), combined with a competition by the other globin genes, are important to downregulate γ-globin gene expression in the adult stage (45,46). Additionally, it is well documented that the distance or gene order relative to the LCR is an essential determinant for the type of globin gene to be activated selectively during the primitive stage of erythropoiesis (47–51). In our study, the LCR was linked directly to the Aγ-globin gene and the lines tested versus the control lines differed solely on either the absence or presence of the Enh and F elements, respectively.

The function of these silencer elements in the context of the entire β-globin cluster, the proper spacing of the LCR, and the high order chromatin structure can be addressed by considering the features of the active chromatin hub (ACH) model (52,53). Specifically, the model refers to the erythroid-specific spatial clustering of *cis* regulatory elements, consisting of the 5' hypersensitive sites HS1-6 of the LCR, the downstream hypersensitive site 3' HS1 and of the active globin genes. According to the ACH model, stable enhancer-promoter interactions determine the establishment of a functional expression module. Depending on the competitive nature of the surrounding chromatin, additional *cis* regulatory elements, such as the Enh and F silencers, may have evolved to stabilize the enhancer-promoter interactions in the ACH and maintain the required expression level at a particular developmental stage. The transcriptional outcome of an ACH critically depends on the productive interaction between *cis* regulatory elements and specific developmental transcription factors, such as Ikaros and its family member Eos, which seem to participate in the regulation of hemoglobin switching by facilitating the DNA looping between the LCR and the upstream region of the δ-globin gene (54). Interestingly, the transcriptional pattern of globin gene switching seems to correlate with the switching of interactions of the individual globin genes with this spatial cluster (55). Furthermore, the ACH model can explain why the γ-globin expression pattern might differ in the case of the Enh and F deletion in the context of the full β-globin locus construct (27), possessing all the regulatory elements of the locus versus the deletion of the same elements in the context of a cosmid construct containing just the LCR and the Aγ gene; it is conceivable that the entry of new regulatory elements may further stabilize or destabilize existing interactions and alter the expression levels of genes present in the ACH. Further investigation on the role of Enh and F elements on γ-globin gene transcription can be accomplished by

applying the chromosome conformation capture carbon copy (5C) technology, which permits the analysis of millions of chromatin interactions in parallel (56), to understand how such elements interact with γ-globin promoters, the LCR, and other genes within the ACH.

All the above studies strongly indicate that apart from the elements located within the Aγ- to δ-globin intergenic region, hemoglobin switching is a consequence of a complex interplay between developmental stage-specific transcription factors which interact both within the promoters of the various genes and the 5' LCR, as well as of specific modifications of the physical location of a gene within the locus (47,57–59). Thus, sequence-specific transcription factors (activators and/or repressors), such as the recently identified developmental stage-specific repressor BCL11A (10), can regulate ε- and γ-globin activation and repression through nearby *cis* elements in a gene autonomous manner, independently of the amplifying effects of the LCR.

The transgenic animal model used in the present study, combined with the data from our previous study (27), allowed us to examine the direct *in vivo* role of these two gene-proximal regulatory elements (Enh and F) in the absence of any additional silencer elements. The main impetus behind this study had been the elucidation of the molecular basis of HPFH as a result of large deletions in the human β-globin locus, which often include the Enh and F elements. Two models have been proposed to account for the generation of the HPFH phenotypes; the deletion of silencer elements in the Aγ to δ-gene region (17), and/or the juxtaposition of downstream enhancers (18–21). The novel data described here and in previous studies (25–27,59) suggest that elements located in the Aγ to δ intergenic region contribute to specific regulation of the individual globin genes at different developmental stages. Taken together, these studies provide a new framework for further investigations for the molecular basis of the phenomenon of human hemoglobin switching.

In addition, our studies employing the specific and selective Enh/F deletion, permitted for the first time the direct *in vivo* evaluation of the effect of the Enh and F elements on the expression of the fetal γ-globin gene in adult life and provide the first experimental proof that these two elements represent *bona fide* repressing elements capable of altering the developmental pattern of expression of the fetal γ-genes. Furthermore, previous studies described by our group (27) and by others (59,60) identified a new class of gene-proximal regulatory elements within the human β-globin locus that are involved in regulating the levels of globin gene transcription either positively (59,60) or negatively (27), in a developmental-stage specific manner.

Further delineation of the molecular mechanisms of fetal-to-adult switch are expected to contribute to designing future therapeutic strategies for thalassemias and sickle cell anemia. To this end, eventual manipulation of Enh, F, and other regulatory elements of the locus or of stage-specific repressors (10) aiming to the derepression and reactivation of the fetal gene expression in the adult erythropoietic stem cells is expected to lead to an effect which has been shown to represent a potent therapeutic approach (7).

DISCLOSURES

The authors declare that they have no competing interests as defined by *Molecular Medicine*, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES

1. Grosveld F, van Assendelft GB, Greaves DR, Kollias G. (1987) Position-independent, high-level expression of the human β-globin gene in transgenic mice. *Cell*. 51:975–85.
2. Ragoczy T, Bender MA, Telling A, Byron R, Groudine M. (2006) The locus control region is required for association of the murine β-globin locus with engaged transcription factories during erythroid maturation. *Genes. Dev.* 20:1447–57.
3. Felsenfeld G, Groudine M. (2003) Controlling the double helix. *Nature*. 421:448–53.
4. Strouboulis J, Dillon N, Grosveld F. (1992) Developmental regulation of a complete 70-kb human β-globin locus in transgenic mice. *Genes. Dev.* 6:1857–64.
5. Townes TM, Behringer RR. (1990) Human globin locus activation region (LAR): role in temporal control. *Trends Genet*. 6:219–23.
6. Brand M, *et al.* (2004) Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat. Struct. Mol. Biol.* 11:73–80.
7. Ley TJ, *et al.* (1982) 5-azacytidine selectively increases γ-globin synthesis in a patient with β+ thalassemia. *N. Engl. J. Med.* 307:1469–75.
8. Gilman JG, Huisman THJ. (1985) DNA sequence variation associated with elevated fetal Gγ globin production. *Blood*. 66:783–7.
9. Creary LE, *et al.* (2009) Genetic variation on chromosome 6 influences F cell levels in healthy individuals of African descent and HbF levels in sickle cell patients. *PLoS ONE*. 4:e4218.
10. Sankaran VG, *et al.* (2008) Human fetal hemoglobin expression is regulated by the developmental stage-specific repressor BCL11A. *Science*. 322:1839–42.
11. Thein SL, Menzel S. (2009) Discovering the genetics underlying foetal haemoglobin production in adults. *Brit. J. Haematol*. 145:455–67.
12. Ragusa A, *et al.* (1992) Genetic epidemiology of β-thalassemia in Sicily: do sequences 5′ to the Gγ gene and 5′ to the β gene interact to enhance HbF expression in β-thalassemia? *Am. J. Hematol.* 40:199–206.
13. Gonçalves I, *et al.* (1998) Combined effect of two different polymorphic sequences within the β globin gene cluster on the level of HbF. *Am. J. Hematol.* 57:269–76.
14. Gonçalves I, *et al.* (2002) Fetal hemoglobin elevation in Hb Lepore heterozygotes and its correlation with β-globin cluster linked determinants. *Am. J. Hematol*. 69:95–102.
15. Bandyopadhyay S, *et al.* (2005) Two β-globin cluster-linked polymorphic loci in thalassemia patients of variable levels of fetal hemoglobin. *Eur. J. Haematol*. 75:47–53.
16. Wood WG. (1993) Increased HbF in adult life. *Baillieres Clin. Haematol.* 6:177–213.
17. Huisman TH, *et al.* (1974) The present status of the heterogeneity of fetal hemoglobin in β-thalassemia: an attempt to unify some observations in thalassemia and related conditions. *Ann. N. Y. Acad. Sci*. 232:107–24.
18. Flavell RA, *et al.* (1983) Structure and expression of the human globin genes and murine histocompatibility antigen genes. *Cold Spring Harb. Symp. Quant. Biol.* 47:1067–78.
19. Anagnou NP, *et al.* (1995) Sequences located 3′ to the breakpoint of the hereditary persistence of fetal hemoglobin-3 deletion exhibit enhancer activity and can modify the developmental expression of the human fetal Aγ-globin gene in transgenic mice. *J. Biol. Chem*. 270:10256–63.
20. Katsantoni EZ, *et al.* (2003) Persistent γ-globin expression in adult transgenic mice is mediated by HPFH-2, HPFH-3, and HPFH-6 breakpoint sequences. *Blood*. 102:3412–9.
21. Arcasoy MO, *et al.* (1997) High levels of human γ-globin gene expression in adult mice carrying a transgene of deletion-type hereditary persistence of fetal hemoglobin. *Mol. Cell Biol.* 17:2076–89.
22. Kosteas T, Manifava M, Moschonas N, Anagnou NP. (1993) Functional analysis of the Aγ to ψβ globin gene region of the human β-locus: evidence for negative regulatory elements. *Clin. Res.* 41:38A.
23. Kosteas T, Manifava M, Moschonas N, Anagnou NP. (1994) Functional analysis of the Aγ to δ globin gene region of the β-cluster: evidence for negative regulatory elements. *Blood*. 84:506A.
24. Dillon N, Grosveld F. (1991) Human γ-globin genes silenced independently of other genes in the β-globin locus. *Nature*. 350:252–4.
25. Gaensler KM, *et al.* (2003) Sequences in the Aγ-δ intergenic region are not required for stage-specific regulation of the human β-globin gene locus. *Proc. Natl. Acad. Sci. U. S. A.* 100:3374–9.
26. Liu Q, Tanimoto K, Bungert J, Engel JD. (1998) The Aγ-globin 3′ element provides no unique function(s) for human β-globin locus gene regulation. *Proc. Natl. Acad. Sci. U. S. A.* 95:9944–9.
27. Katsantoni EZ, *et al.* (2004) An embryonic-specific repressor element located 3′ to the Aγ-globin gene influences transcription of the human β-globin locus in transgenic mice. *Exp. Hematol.* 32:224–33.
28. Dillon N, Grosveld F. (1993) Transcriptional regulation of multigene loci: multilevel control. *Trends Genet.* 9:134–7.
29. Sakai K, Miyazaki J. (1997) A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochem. Biophys. Res. Commun.* 237:318–24.
30. Milot E, *et al.* (1996) Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell*. 87:105–14.
31. Lindenbaum MH, Grosveld F. (1990) An *in vitro* globin gene switching model based on differentiated embryonic stem cells. *Genes. Dev.* 4:2075–85.
32. Fraser P, Hurst J, Collis P, Grosveld F. (1990) DNaseI hypersensitive sites 1, 2 and 3 of the human β-globin dominant control region direct position-independent expression. *Nucleic Acids Res.* 18:3503–8.
33. Pfaffl MW. (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45.
34. Mulder MP, *et al.* (1995) Positional mapping of

loci in the DiGeorge critical region at chromosome 22q11 using a new marker (D22S183). *Hum. Genet*. 96:133–41.

35. Schmidt M, *et al*. (2002) Polyclonal long-term repopulating stem cell clones in a primate model. *Blood*. 100:2737–43.

36. Yamada Y, *et al*. (2003) Neuropilin-1 on hematopoietic cells as a source of vascular development. *Blood*. 101:1801–9.

37. Kollias G, Wrighton N, Hurst J, Grosveld F. (1986) Regulated expression of human Aγ-, β-, and hybrid γβ-globin genes in transgenic mice: manipulation of the developmental expression patterns. *Cell*. 46:89–94.

38. Collis P, Antoniou M, Grosveld F. (1990) Definition of the minimal requirements within the human β-globin gene and the dominant control region for high level expression. *EMBO J*. 9:233–40.

39. Enver T, Ebens AJ, Forrester WC, Stamatoyannopoulos G. (1989) The human β-globin locus activation region alters the developmental fate of a human fetal globin gene in transgenic mice. *Proc. Natl. Acad. Sci. U. S. A*. 86:7033–7.

40. Lloyd JA, Krakowsky JM, Crable SC, Lingrel JB. (1992) Human γ- to β-globin gene switching using a mini construct in transgenic mice. *Mol. Cell Biol*. 12:1561–7.

41. Bodine DM, Ley TJ. (1987) An enhancer element lies 3′ to the human Aγ globin gene. *EMBO J*. 6:2997–3004.

42. Purucker M, Bodine D, Lin H, McDonagh K, Nienhuis AW. (1990) Structure and function of the enhancer 3′ to the human Aγ globin gene. *Nucleic Acids Res*. 18:7407–15.

43. Chakalova L, *et al*. (2005) The Corfu δβ thalassemia deletion disrupts γ-globin gene silencing and reveals post-transcriptional regulation of HbF expression. *Blood*. 105:2154–60.

44. Anagnou NP, Karlsson S, Moulton AD, Keller G, Nienhuis AW. (1986) Promoter sequences required for function of the human γ-globin gene in erythroid cells. *EMBO J*. 5:121–6.

45. Ragoczy T, Telling A, Sawado T, Groudine M, Kosak ST. (2002) A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. *Chromosome Res*. 11:513–25.

46. Omori A, Tanabe O, Engel JD, Fukamizu A, Tanimoto K. (2005) Adult stage γ-globin silencing is mediated by a promoter direct repeat element. *Mol. Cell Biol*. 25:3443–51.

47. Harju S, Navas PA, Stamatoyannopoulos G, Peterson KR. (2005) Genome architecture of the human β-globin locus affects developmental regulation of gene expression. *Mol. Cell Biol*. 25:8765–78.

48. Dillon N, Trimborn T, Strouboulis J, Fraser P, Grosveld F. (1997) The effect of distance on long-range chromatin interactions. *Mol. Cell*. 1:131–9.

49. Hanscombe O, *et al*. (1991) Importance of globin gene order for correct developmental expression. *Genes. Dev*. 5:1387–94.

50. Peterson KR, *et al*. (1993) Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human β-globin locus display proper developmental control of human globin genes. *Proc. Natl. Acad. Sci. U. S. A*. 90:7593–7.

51. Tanimoto K, Liu Q, Bungert J, Engel JD. (1999) Effects of altered gene order or orientation of the locus control region on human β-globin gene expression in mice. *Nature*. 398:344–8.

52. Noordermeer D, de Laat W. (2008) Joining the loops: β-globin gene regulation. *IUBMB Life*. 60:824–33.

53. Palstra R-J, *et al*. (2003) The β-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet*. 35:190–4.

54. Keys JR, *et al*. (2008) A mechanism for Ikaros regulation of human globin gene switching. *Brit. J. Haematol*. 141:398–406.

55. Stamatoyannopoulos G, Grosveld F. (2001) Hemoglobin switching. In: *The Molecular Basis of Blood Diseases*. Stamatoyannopoulos G, Majerus P, Perlmutter R, Varmus H (eds.) W.B. Saunders, Philadelphia, pp. 135–82.

56. van Bekkum NL, Dekker J. (2009) Determining spatial chromatin organization of large genomic regions using 5C technology. *Methods Mol. Biol*. 567:189–213.

57. Raich N, Clegg CH, Grofti J, Roméo PH, Stamatoyannopoulos G. (1995) GATA1 and YY1 are developmental repressors of the human ε-globin gene. *EMBO J*. 14:801–9.

58. Rupon JW, Wang SZ, Gaensler K, Lloyd J, Ginder GD. (2006) Methyl binding domain protein 2 mediates γ-globin gene silencing in adult human βYAC transgenic mice. *Proc. Natl. Acad. Sci. U. S. A*. 103:6617–22.

59. Liu Q, Bungert J, Engel JD. (1997) Mutation of gene-proximal regulatory elements disrupts human ε-, γ-, and β-globin expression in yeast artificial chromosome transgenic mice. *Proc. Natl. Acad. Sci. U. S. A*. 94:169–74.

60. Calzolari R, McMorrow T, Yannoutsos N, Langeveld A, Grosveld F. (1999) Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and δβ-thalassemia affects β- but not γ-globin gene expression. *EMBO J*. 18:949–58.