4-23-2007

CpG Hypomethylation in a Large Domain Encompassing the Embryonic β-Like Globin Genes in Primitive Erythrocytes

Mei Hsu
Dartmouth College

Rodwell R. Mabaera
Dartmouth College

Christopher H. Lowrey
Dartmouth College

David I. K. Martin
Children's Hospital Oakland Research Institute

Steven Fiering
Dartmouth College

MOLECULAR AND CELLULAR BIOLOGY, July 2007, p. 5047–5054 Vol. 27, No. 13
0270-7306/07/$08.00+0 doi:10.1128/MCB.02234-06

CpG Hypomethylation in a Large Domain Encompassing the Embryonic β-Like Globin Genes in Primitive Erythrocytes▽†

Mei Hsu,[1] Rodwell Mabaera,[2] Christopher H. Lowrey,[2] David I. K. Martin,[3] and Steven Fiering[1]*

Department of Microbiology and Immunology, Department of Genetics, and Norris Cotton Cancer Center, Dartmouth Medical School, Hanover, New Hampshire[1]; Department of Medicine, Department of Pharmacology and Toxicology, and Norris Cotton Cancer Center, Dartmouth Medical School, Hanover, New Hampshire[2]; and Children's Hospital Oakland Research Institute, Oakland, California[3]

Received 28 November 2006/Returned for modification 22 January 2007/Accepted 12 April 2007

There is little evidence addressing the role of CpG methylation in transcriptional control of genes that do not contain CpG islands. This is reflected in the ongoing debate about whether CpG methylation merely suppresses retroelements or if it also plays a role in developmental and tissue-specific gene regulation. The genes of the β-globin locus are an important model of mammalian developmental gene regulation and do not contain CpG islands. We have analyzed the methylation status of regions in the murine β-like globin locus in uncultured primitive and definitive erythroblasts and other cultured primary and transformed cell types. A large (~20-kb) domain is hypomethylated only in primitive erythroid cells; it extends from the region just past the locus control region to before β-major and encompasses the embryonic genes Ey, βh1, and βh0. Even retrotransposons in this region are hypomethylated in primitive erythroid cells. The existence of this large developmentally regulated domain of hypomethylation supports a mechanistic role for DNA methylation in developmental regulation of globin genes.

Chromatin structure is a central control point in eukaryotic transcriptional regulation; it determines where transcription is initiated and what DNA is transcribed. Transcription initiation is stably suppressed in large portions of the genome: retroelements are constitutively suppressed, and genes are suppressed in tissue-specific patterns by very similar mechanisms. The chromatin makeup of a region of DNA, termed its "epigenotype," is a complex assortment of proteins and chemical modifications that differs between and even within specific cell types. Cytosine methylation, histone modification, histone variants, and nonhistone chromatin proteins have interacting roles in relatively stable chromatin structures. Dense cytosine methylation is associated with suppression of transcription initiation (reviewed in reference 16), and it appears to inhibit but not prevent transcription once initiated (20).

In mammals, DNA methylation occurs predominantly at CpG dinucleotides. Clusters of CpGs are rare in the mammalian genome outside of CpG islands, which are defined as regions of generally 200 to 500 bp with an unusually high frequency of CpGs. Many CpG islands are associated with gene promoters (30). Hypermethylation of CpG islands is rare in untransformed cells and more common in transformed cells; regardless of cell type, it is associated with transcriptional suppression of the methylated gene. The relationship between CpG methylation and expression of genes without CpG island promoters, such as those at the β-globin locus, is less clear.

Outside of CpG islands, the frequency of CpG dinucleotides is lower than expected from genomic CG content, and roughly 80% of CpG sites outside of CpG islands are methylated (5). There have been no reports of large (multikilobase) regions of hypomethylated CpG dinucleotides in the mammalian genome. A large proportion of CpG methylation occurs at retroelements that are invariably hypermethylated (35). Although it has been supposed that CpG methylation is also associated with developmental and tissue-specific silencing of genes, there are limited examples of developmental changes in methylation patterns at specific genes (7, 10).

The mammalian β-globin locus has been studied as a model of developmentally regulated transcription in a multigene locus; developmental globin "switching" is also of great interest, because a ready means of reversing the fetal-to-adult switch in the human β-globin locus would provide a treatment for the devastating hemoglobinopathies β-thalassemia and sickle cell disease. Like other mammalian β-globin loci, the murine β-globin locus is so structured that genes share a common transcriptional orientation and are arranged roughly in their developmental order of expression. The developmental switch in globin gene transcription occurs concomitantly with changes in the anatomic site of erythropoiesis and in the type of erythrocyte (Fig. 1). In mice and most other mammals (excluding humans and other higher primates), there is a single switch that occurs when primitive erythropoiesis in the embryonic yolk sac is replaced by definitive erythropoiesis in the fetal liver. Nucleated primitive red cells produced in the yolk sac of murine embryos express high levels of the embryonic globins εy and βH1 and very low levels of the fetal/adult-stage globins, β-major and β-minor (27). In fetal liver-derived definitive erythroblasts, β-major and β-minor are expressed at high levels and the embryonic globins are silent, a condition maintained into adulthood after erythropoiesis shifts to the bone marrow, where definitive erythrocytes are produced throughout postnatal life. Unlike humans, who switch from γ to β when eryth-

* Corresponding author. Mailing address: Dartmouth Medical School, 6 Rubin, Dartmouth Hitchcock Medical Center, Lebanon, NH 03756. Phone: (603) 653-9966. Fax: (603) 653-9952. E-mail: fiering@dartmouth.edu.

† Supplemental material for this article may be found at http://mcb.asm.org/.

▽ Published ahead of print on 23 April 2007.

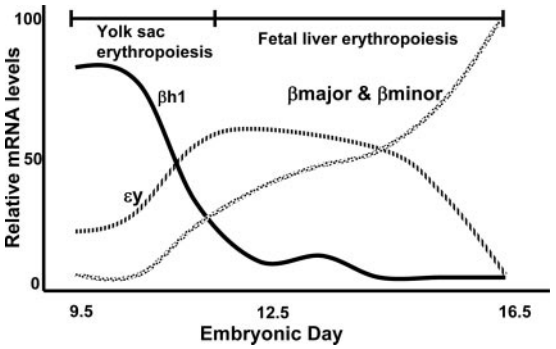


FIG. 1. Changes in levels of β-like globin mRNAs during development. Note that the embryonic genes εy and βh1 are expressed in primitive erythroid cells generated in the yolk sac during early embryonic development, with transcripts undetectable by E16.5. Data are adapted from reference 33.

ropoiesis moves from fetal liver to bone marrow, there are no adult-specific β-like globin genes expressed in murine bone marrow. However, there is a shift in the ratio of β-major to β-minor expression between fetal liver-derived (60% β-major, 40% β-minor) and bone marrow-derived (80% β-major, 20% β-minor) definitive erythrocytes (34).

Several lines of evidence have long suggested that cytosine methylation plays a role in β-like globin regulation, but a clear picture of its function is still lacking. Drugs capable of modifying DNA methylation can induce γ-globin expression in baboons, in transgenic mice carrying the human β-globin locus YAC, and in patients with hemoglobinopathies (2, 3, 17, 19, 26). Mutation of MBD2, a protein that binds regions of CpG methylation and suppresses transcription, increases expression of γ-globin from the human β-globin locus YAC in transgenic mice (29). Southern blot analysis of DNA digested with methylation-sensitive restriction enzymes showed a general correlation between DNA hypomethylation in or close to the promoter of an active β-like globin gene and its transcription in primary human erythroblasts (21, 25, 32); however, a very limited number of CpG dinucleotides could be assessed by this method, leaving the picture incomplete.

We have carried out a detailed analysis of the patterns of CpG methylation associated with developmental regulation of the murine β-like globin locus, using bisulfite allelic sequencing of DNA from primitive and definitive mouse erythroblasts. The study reveals a surprisingly large hypomethylated domain that surrounds the expressed embryonic genes in primitive cells. Retroelements that are found within this domain are also hypomethylated in primitive erythrocytes. This adds to mounting evidence that the primitive erythrocytes regulate globins very differently than definitive erythrocytes. A variety of studies now suggest that primitive erythrocytes are a fundamentally different lineage than definitive erythrocytes, and globin regulatory mechanisms may reflect this lineage divergence (reviewed in reference 22). These findings demonstrate a pattern of CpG methylation in the murine β-globin locus that is consistent with a mechanistic role in the developmentally regulated gene expression program.

MATERIALS AND METHODS

Cells. Circulating primitive red blood cells and whole fetal livers were collected from embryonic day 11.5 (E11.5) and E14.5 CD-1 strain mouse embryos, respectively. E11.5 circulating cells are virtually pure primitive erythrocytes. Whole fetal liver is reported to be 80 to 92% erythroid cells (14). This was confirmed by flow cytometry using anti-Ter-119-stained cells. Erythroblasts from fetal liver were either not further purified or purified with anti-Ter-119 magnetic microbeads (Miltenyi Corp., Auburn, CA) (see Fig. 3A, below). Erythroblasts from all adult bone marrow samples (and some fetal livers) were purified with anti-Ter-119 magnetic microbeads to ≥94% purity, as determined by flow cytometry. Animal experimentation was under the auspices of a protocol approved by the Dartmouth IACUC and was compliant with all relevant regulations of the NIH Office of Laboratory Animal Welfare.

R1 embryonic stem (ES) cells were grown on mouse embryonic fibroblast (MEF) feeders in Dulbecco's modified Eagle's medium supplemented with 15% fetal bovine serum, 1,000 units/ml ESGRO (Chemicon International, Termecula, CA), 1% L-glutamine, 1% penicillin-streptomycin, and 14 mM β-mercaptoethanol. MEFs derived from E14.5 CD-1 embryos and mouse erythroleukemia (MEL) cells were grown in Dulbecco's modified Eagle's medium with 10% fetal bovine serum, 1% L-glutamine, and 1% penicillin-streptomycin.

Bisulfite conversion was performed as described elsewhere (28). In brief, 2 to 3 μg of genomic DNA in 20 μl of water was denatured with a 20-min incubation at 37°C after the addition of 2.2 μl 3 M NaOH. A 208-μl volume of 6.24 M urea–4 M bisulfite and 12 μl of 10 mM hydroquinone were added prior to incubating for 4 h at 55°C. The reaction mixture was desalted with Promega Wizard DNA CleanUp columns per the manufacturer's instructions. DNA was desulfonated by addition of a 1/10 volume of 3 M NaOH and left at 37°C for 20 min, followed by ethanol precipitation. Sequences of interest were amplified by nested or seminested PCR (see Table S2 in the supplemental material) with 200 μM deoxynucleoside triphosphates, 100 μM of each primer, 3 mM $MgCl_2$, 50 mM KCl, 0.01% Triton X-100, and 10 mM Tris, pH 9, and cycled at 94°C for 2 min, five times at 94°C for 1 min, 50°C for 2 min, and 72°C for 1 min 30 s, 35 times at 94°C for 30 s, 50°C for 2 min, and 72°C for 1 min 30 s, and finally at 72°C for 6 min. PCR products were gel purified and TA cloned with the Promega T-easy kit. Plasmids were purified with a QiaQuick spin miniprep kit and submitted to the Dartmouth College Molecular Biology Core Facility for sequencing.

Reverse transcription-PCR (RT-PCR) studies. cDNA was reversed transcribed from total cellular RNA isolated from circulating cells of E10.5 yolk sacs and Ter-119-positive bone marrow cells with QIAGEN RNeasy columns with additional DNase treatment. Template concentrations were equalized using primers to murine β-actin. PCR was carried out in a 20-μl volume with 2 U of *Taq*, 15 mM $MgCl_2$, 1 μM primer, and 200 μM deoxynucleoside triphosphates. After incubating at 95°C for 2 min, reactions were amplified for 35 cycles (95°C for 45 s, 58 to 62°C for 1 min, and 72°C for 45 s). Products were separated on a 2% agarose gel stained with ethidium bromide and recorded with a Bio-Rad GelDoc photo documentation system.

RESULTS

Our general strategy for studying the DNA methylation status of the globin locus is to isolate cells of interest and perform bisulfite allelic sequencing on regions of interest within the globin locus. Allelic sequencing allows one to assess the pattern of CpG methylation in a given region: when sufficient numbers of alleles are sequenced, both average density and extent of variation between alleles are readily apparent. By sampling multiple alleles from each of several regions, we assembled a picture of methylation state in the region as a whole. For convenience of presentation, hypermethylation is defined as 67% or more methylated CpGs in the total of all CpGs in all alleles of an amplicon; intermediate methylation is defined as 33 to 66% methylated; hypomethylation is defined as 32% or less.

Promoter methylation status. Because CpG methylation near the start site of transcription is most likely to reflect transcription state and/or potential, we initially analyzed methylation at the promoters of the εy, βh1, β-major, and β-minor genes in ES cells, primitive erythroblasts from embryonic cir-

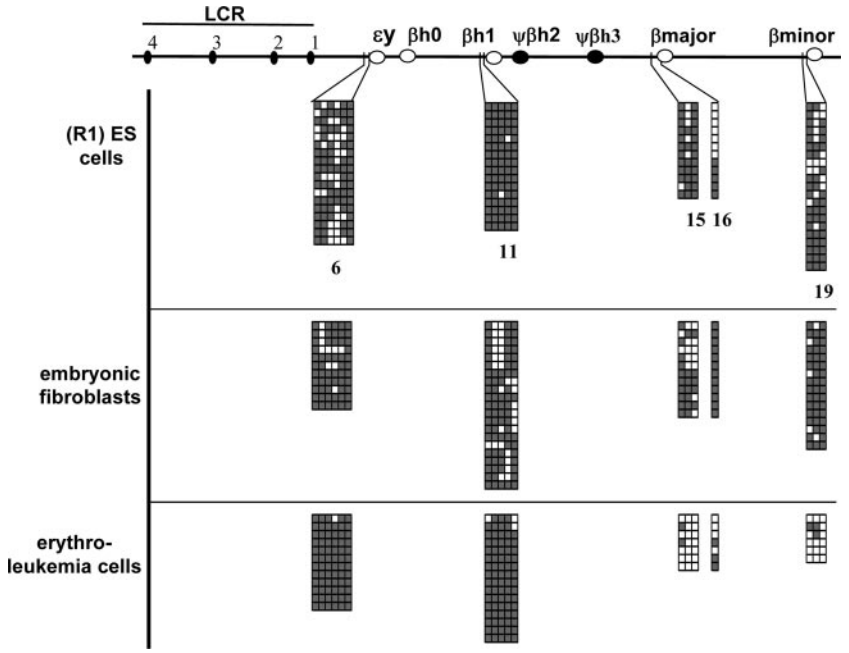


FIG. 2. CpG methylation of murine β-like globin genes in cultured cells. Top, scaled map of the murine β-like globin locus. Vertical lines demarcate segments assayed. Open ovals, transcribed genes; filled ovals, pseudogenes; numbered filled ovals, DNase-hypersensitive sites in the LCR. Below, each horizontal row of blocks represents one bisulfite-treated clone of the indicated segment; vertical columns denote specific CpG sites. Open box, unmethylated; filled box, methylated.

culation, definitive erythrocytes from the fetal liver and bone marrow, MEL cells, and MEF cells. This provides a developmental range of cells that either express β-like globins (primitive and definitive erythroblasts and MEL), can differentiate into globin-expressing cells (ES), or never express any globin genes (MEFs). All are primary cells except for MEL cells. Each promoter was studied across 300 to 600 bp around the transcriptional start site (Fig. 2, segments 6, 11, 15, 16, and 19).

ES and MEF cells do not express β-like globins, and in these cells all four of the promoters were hypermethylated (Fig. 2). In MEL cells, which can express high levels of β-major and β-minor when induced to differentiate, the promoters of both genes were hypomethylated, while the promoters of the embryonic genes εy and βh1 (which are not expressed in MEL cells) were hypermethylated.

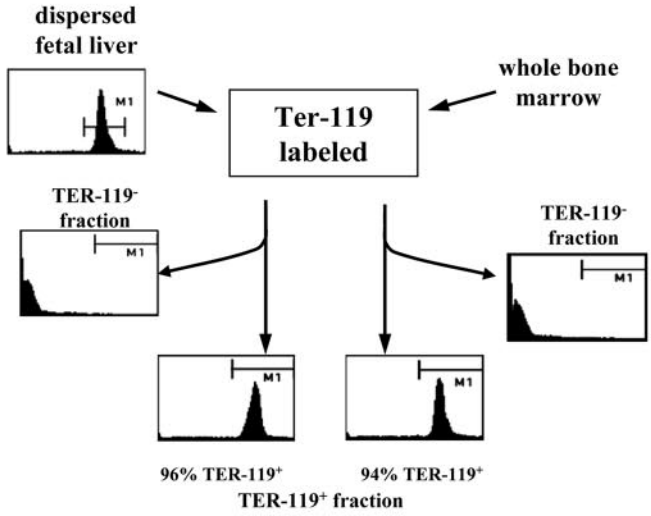
Primary primitive erythroid cells were isolated from the circulation of E11.5 embryos. Definitive erythroid cells from fetal livers (E14.5) or adult bone marrow were generally purified by selection for the pan-erythroid Ter-119 antigen, and purity was validated by flow cytometry analysis for Ter-119 expression (Fig. 3A). In primary erythroid cells, hypomethylation of the embryonic gene promoters is fully correlated with expression (Fig. 3B). The εy (segment 6) and βH1 (segment 11) promoters are hypomethylated (4 to 8% methylation) in yolk sac-derived circulating primitive red cells and hypermethylated in definitive red cell precursors from either fetal liver or bone marrow (εy, >71%; βH1, >68%). Surprisingly the β-major and β-minor promoters do not follow this simple pattern. In primitive cells where both are expressed at relatively low levels, the β-major promoter (segment 15) has 26% methylation and the β-minor promoter (segment 19) is 58% methylated. Inspection of the allelic pattern, however, shows that for the β-major promoter in particular, there is high interallelic variation, with some alleles fully methylated and others completely unmethylated. In definitive cells where both are highly expressed, the β-major promoter is hypomethylated (less than 5% methylated) and the β-minor promoter continues to be moderately methylated (33% to 47% methylated). Despite hypomethylation of the β-major promoter in definitive cells, hypomethylation does not extend into the 5′ untranslated region of the transcribed portion of β-major. The CpGs at positions +114 and +117 relative to the β-major transcriptional start (segment 16) are amplified along with the promoter and are moderately methylated (fetal liver, 35%) or hypermethylated (bone marrow, 87%).

Methylation patterns of nonrepetitive sequence outside the promoters reveal a domain of hypomethylation. The consistency of promoter hypomethylation in the embryonic gene promoters in primitive cells and the variability of promoter methylation of the adult genes in both primitive and definitive cells suggested that the embryonic and fetal/adult globin genes could also have different methylation states associated with nonpromoter sequences. Therefore, we expanded the methylation survey of primary erythroid cells to include several nonpromoter regions: the εy coding region (segment 7), the coding region of the weakly expressed embryonic gene βH0 (segment 9), the region between the βh0 and βh1 embryonic genes (segment 10), the region 5′ of βH2 (segment 12; an unexpressed gene homologous to the embryonic globin genes) (9), the intergenic region upstream of β-major (segment 14), the β-major coding region (segment 17), and the region just 3′ to β-major between β-major and β-minor (segment 18) (Fig. 3B).

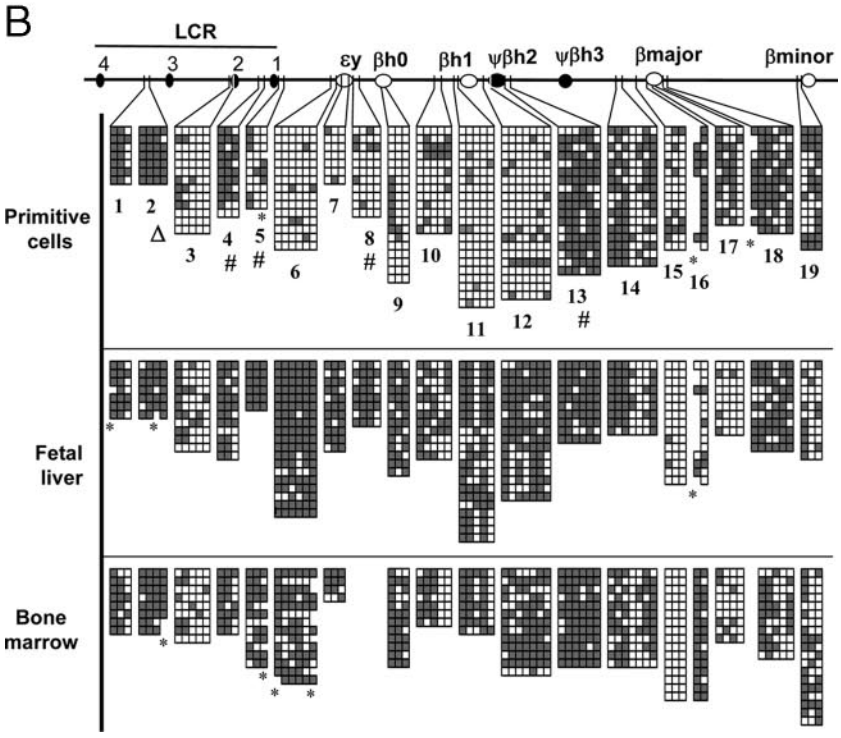
This extended survey of CpG methylation across the murine β-globin locus reveals a domain pattern of methylation, with

A

Cell/Tissue	Segment Assayed	Method of Isolation	Qualitation	Erythroblast Purity
primitive red blood cells	**1-19**	circulating red blood cells from yolk sacs of E11.5 embryos		
fetal liver definitive erythroblasts	**6, 11, 12**	whole fetal liver from E14.5 embryos	FACscan	80-92%
	1-5, 7-10, 13-14, 16-17	positively selected TER-119 expressing cells	FACscan and/or slide	98-99.8%
	15	whole liver and TER-119 cells	FACscan or slide	80-99.8%
adult (bone marrow) definitive erythroblasts	**1-13, 16-19**	positively selected TER-119 expressing cells	slide	80-92%
	14-15	positively selected TER-119 expressing cells	FACscan & slide	93-97%



B



developmental variation. Primitive red cells exhibit a strikingly different methylation pattern than definitive red cells, while there are minimal differences between definitive cells from fetal liver and bone marrow (Fig. 3B). In primitive red cells, which predominantly express εy and βh1 (βh0, β-major, and β-minor are expressed at lower levels), not only the promoters but also the intragenic and surrounding intergenic regions of the embryonic globins are hypomethylated (segments 7, 9, 10, and 12). Sequences flanking and within the embryonic globins, εy and βh1, are all hypermethylated in definitive erythroblasts. In contrast, the intergenic regions surrounding β-major are moderately methylated or hypermethylated in all primary erythroid cells (segments 14 and 18). In definitive cells, the gene body of β-major is variably methylated, with the region near the 5′ end of the gene moderately methylated or hypermethylated (segment 16) and the region near the 3′ end hypomethylated (segment 17).

Repetitive elements in the hypomethylated domain are also hypomethylated. These results clearly delineate a previously unreported DNA methylation pattern: a large domain (~20 kb) in which CpG dinucleotides in nonrepetitive sequences are hypomethylated in a developmentally regulated pattern. To further probe this novel hypomethylated domain, we analyzed the methylation status of repetitive elements around the embryonic genes using a PCR primer within the repetitive element and a primer in the unique sequence flanking the element. In embryonic erythroblasts, a LINE element between εy and βh1 (segment 8) is hypomethylated in primitive erythroblasts (16%) but hypermethylated in definitive erythroblasts (75%); a similar pattern is seen in a LINE element upstream of εy (segment 5). However, a similar LINE element between βh1 and β-major (segment 13) (Fig. 2B) is hypermethylated in both primitive (76%) and definitive (90%) erythroblasts and appears to form the 3′ boundary of this hypomethylated domain in primitive erythroblasts.

The LCR is not part of the hypomethylated domain. The locus control region (LCR) is a group of transcriptional enhancers that together form an important regulatory element in the β-globin locus; it lies upstream of the embryonic hypomethylated region delineated by these studies. In order to determine whether the LCR is hypomethylated and to define the 5′ boundary of the hypomethylated region, we assessed the methylation status in primitive and definitive cells of hypersensitive site 2 (HS2) of the LCR as well as repetitive and unique elements around HS1, -2, and -3. HS2 was hypomethylated or moderately methylated in all samples (segment 3). Repetitive elements within the LCR were hypermethylated in both primitive and definitive erythroblasts (segments 2 and 4), and the unique segment 1, which is within the LCR but not a hypersensitive site, was hypermethylated in all samples. This result indicates that the LCR is not part of the embryonic hypomethylated domain and contains the 5′ boundary of this domain at or around HS1 of the LCR.

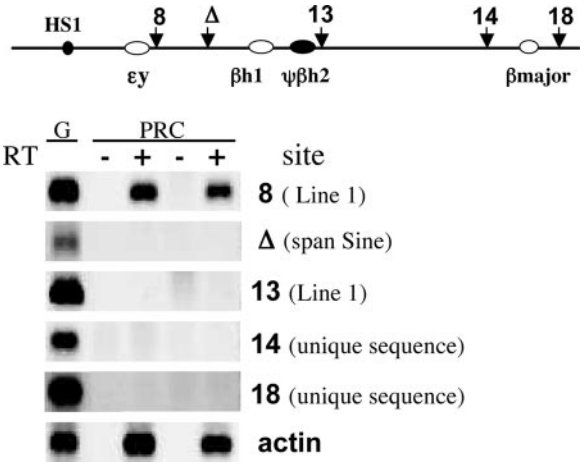


FIG. 4. Nongenic transcription within the hypomethylated domain. The map shows the sites of primers used to assay by RT-PCR. Numbered sites are at regions noted in Fig. 3B. G is a lane with genomic DNA; the + or − RT lanes denote whether reverse transcriptase was in the reaction mixture to produce cDNA and control for contamination with genomic DNA.

Nongenic transcription within the hypomethylated domain. It has been shown that hypomethylation of repetitive elements is associated with their transcriptional activation (11, 12). This raises the possibility that the repetitive elements in the primitive cell hypomethylated domain are transcribed. RT-PCR of unfractionated RNA was utilized to investigate this possibility (Fig. 4). The LINE 1 element that is just downstream of εy (site 8) is clearly transcribed in the region known to be hypomethylated but a short Sine element is not, and the LINE 1 element that is hypermethylated and appears to be the boundary of the region of the hypomethylated domain is not transcribed in primitive cells. The same sites were assayed in definitive cells from bone marrow, but no evidence for nongenic transcription was seen (data not shown).

DISCUSSION

We find that hypomethylation of CpGs defines a large domain within the multigenic β-globin locus that encompasses the embryonic genes in primitive erythroid cells. Outside of this domain, CpG hypomethylation is consistently found only at the promoter of the β-major gene in primitive or definitive erythroid cells and in HS2 and possibly other hypersensitive sites of the LCR. Remarkably, the hypomethylated domain includes repetitive elements that are members of families present in thousands of copies dispersed throughout the genome; these retroelements have been generally assumed to be hypermethylated in all genomic contexts in normal primary cells.

Transcription alone is not sufficient to determine hypo-

FIG. 3. Methylation status of CpGs surveyed in primary erythroid cells. (A) Top: isolation method and erythroid purity of samples. Bottom: representative FACScan for Ter-119 expression on Ter-119-sorted cells. (B) Organization of data similar to that in Fig. 2. Assayed segments are numbered for reference in the text. Sites marked with a # are Line elements, and the site with a Δ is a Sine element. The asterisks denote polymorphic CpGs, and unboxed sites in the column did not have that CpG.

methylation of the promoter or transcribed region. Although transcription of β-major and β-minor is a minority of total β-like globin transcription in primitive cells, they generate roughly 5% of the total β-like globin transcription and, therefore, they clearly are transcribed genes in primitive cells. Despite this transcription, the β-minor promoter is hypermethylated in primitive cells and the gene body of β-major is hypermethylated in primitive cells. In definitive cells, where β-major and β-minor are among the most highly transcribed genes, both the β-minor promoter and the β-major gene body have intermediate levels of methylation. In association with the hypomethylation of extragenic elements in the primitive cells, the intermediate to high levels of methylation of the β-minor promoters and β-major transcribed regions suggest that the hypomethylation of CpGs in the embryonic domain is part of the overall transcriptional regulation of the embryonic genes, rather than the result of transcription. These results support the possibility that CpG methylation plays a role in developmental globin switching.

The methylation status of the LCR has not been previously assayed in any system. Our limited analysis indicates that at least HS2 and possibly other HSs are hypomethylated in either primitive or definitive cells, but the regions between the hypersensitive sites, whether unique sequences or repetitive elements, are hypermethylated in both primitive and definitive cells. Clearly, the LCR is not part of the hypomethylated domain around the embryonic region in primitive cells, and if it plays a role in establishing that hypomethylated domain, it is not simply as a source from which hypomethylation can spread into the rest of the locus.

The role of CpG methylation in normal processes of developmental gene regulation remains in doubt. There is no question that hypermethylation is associated with the silent transcriptional state; the molecular basis for this was demonstrated when CpG methylation was shown to recruit histone deacetylase activity (13, 24). Experimental reduction of methylation is associated with transcriptional activation of retroelements (8, 11, 12), demonstrating that one role of methylation is to suppress their transcription. In cancer cells, CpG islands may become hypermethylated, but this most likely represents a pathological extension to cellular genes of the mechanisms that repress retrotransposons. There is, however, some evidence that CpG methylation is part of the normal regulatory apparatus of the maspin gene, which contains a CpG island (4, 7), and the interleukin-4 gene, which does not have a CpG island (10). Our findings suggest that CpG methylation participates in the regulation of an entire domain containing multiple genes and repeat elements, even in the absence of CpG islands anywhere in the multigenic β-globin locus.

We observed hypomethylation of retrotransposon elements in the embryonic domain of the β-like globin locus. Retrotransposons are genomic parasites that typically remain methylated in somatic cells; this state is maintained even when the retrotransposon is part of a transcribed region (D. I. K. Martin, unpublished data). It has been proposed that the primary role of CpG methylation is to suppress the activity of retroelements (35) that can create both genetic and epigenetic instabilities when active. It is thus remarkable that multiple retrotransposons are hypomethylated as part of the embryonic β-like globin domain; the developmentally regulated demethylation of these elements suggests that they have been functionally incorporated into the domain and is further evidence that CpG methylation is part of the regulatory apparatus in this domain. Our data (Fig. 4) show that there is some intergenic RNA generated that includes a hypomethylated LINE 1 element. This RNA is generated only in primitive cells and is not seen in definitive cells, associating the production of this RNA with hypomethylation but not establishing a causal relationship between hypomethylation and transcription of this element.

The repetitive elements analyzed in and around the hypomethylated domain are highly similar in sequence to other members of their respective families, members that are highly methylated when found elsewhere in the genome. This point is illustrated by the LINE elements in segments 4 and 5 that sit on either side of the 5′ boundary of the hypomethylated domain in primitive cells. These elements are extremely similar: both are members of the L1_Mm family, contain the same part of that element (ORF2), are roughly the same length (1,643 or 1,786 bases), and are 92% or 93% identical to the family consensus and 92% identical to each other over 1 kb. Yet, one is methylated and the other is not. This finding indicates that the sequence structure of the repetitive elements that are hypomethylated in the primitive domain is unlikely to explain their unusual epigenetic state.

Several studies have examined the relationship between histone modifications and gene expression in the murine β-like globin locus in primitive and definitive red cells (1, 6, 15). The results of these studies are similar to those we report here, if CpG hypomethylation is equated with those histone modifications associated with active transcription. The region of the β-like globin locus spanning the embryonic genes εy and βH1 to just downstream of the unexpressed βH3 pseudogene is enriched in acetylated histone H3 and in H4 and H3 methylated at lysine 4 in primitive cells (15). The same region is depleted of such modifications in definitive red cells (1); this is similar to the domain of DNA hypomethylation we observed proximal to the embryonic genes in primitive cells. As with DNA hypomethylation, there is no broad region of expression-associated histone modifications around the fetal/adult genes in definitive cell precursors. The promoters, coding regions, and regions very close to β-major and β-minor are enriched in acetylated H3 and methylated H3 at lysine 4 in both primitive and definitive cells, but most of the region between these two genes is not enriched for these epigenetic markers of active genes (1, 6, 15).

It has been difficult to determine if a given chromatin modification directly (mechanistically) affects transcription or, alternatively, if the modified chromatin structure is a consequence of other changes. Our studies have the same limitation. While it is possible that the process of transcription causes changes in chromatin structure, transcription itself neither results in nor requires hypomethylation (although transcription initiation is inhibited by CpG methylation and associated modifications). Active transcription of coding regions and downstream sequences may create or help to maintain the acetylation state of local histones, but the large DNA hypomethylated domain in primitive erythrocytes extends well beyond the genes and so cannot be attributed to transcription of the genes themselves.

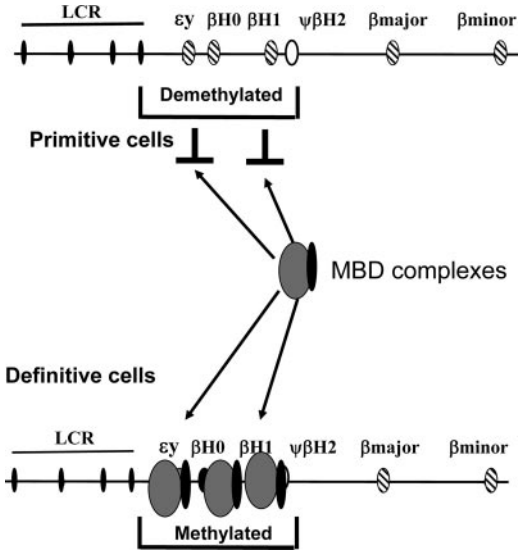


FIG. 5. Model of interaction of methyl binding domain complexes around embryonic genes. Top, in primitive cells the lack of methylation across the domain blocks recruitment of MBD complexes. Bottom, methylation of the domain in definitive cells permits recruitment of MBD complexes, which contribute to gene silencing of the embryonic genes. Expressed genes are shown as hatched ovals.

Most studies of mammalian DNA methylation have focused on CpG island genes, which in comparison to the rest of the genome have a high density of CpG dinucleotides. The mammalian β-like globin loci do not contain any CpG islands. The influence of low-density CpG methylation on transcription remains uncertain, but it is clear that transcription of at least some non-CpG island genes is influenced by DNA methylation. The interleukin-4 gene, which like the β-globin locus has a low density of CpG dinucleotides, is regulated at least in part by the MBD2 methyl-DNA binding protein (10). MeCP2 will bind to a single methylated CpG in vitro (18, 23), providing a mechanistic basis for the view that even isolated CpGs can affect epigenetic regulation through recruitment of methylcytosine binding proteins, which can themselves recruit other proteins, like histone deacetylases. In addition, repetitive elements, which clearly are suppressed by methylation, do not usually contain CpG islands. We suggest that a limited number of methylated CpGs over a broad region could recruit methyl-C binding proteins and thereby participate in the silencing of the embryonic genes in definitive cells, as diagrammed in Fig. 5.

It has been proposed that the transition from fractional to global genome methylation occurred close to or at the origin of the vertebrate subphylum (31). Given this, it is interesting to speculate that the hypomethylated domain observed over the embryonic globins described within predates the origin of vertebrates. This idea is supported by the concept that primitive red cells are the original lineage of red cells and have been supplanted by definitive red cells: cells of the primitive type are nowhere observed in adult vertebrates, but we speculate that they were used exclusively by some progenitor of this subphylum.

ACKNOWLEDGMENTS

We thank Eric Bouhassira for technical advice, Ron Kinser for technical help, and the Dartmouth College Molecular Biology Core for sequencing.

This work was supported by Public Health Service grants DK54071 from NIH NIDDK and HL73431 from NIH NHLBI (to S.F.).

REFERENCES

1. **Bulger, M., D. Schubeler, M. A. Bender, J. Hamilton, C. Farrell, M. R. C. Hardison, and M. Groudine.** 2003. A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse β-globin locus. Mol. Cell. Biol. **23:**5234–5244.
2. **Charache, S., G. Dover, K. Smith, C. C. Talbot, M. Moyer, and S. Boyer.** 1983. Treatment of sickle cell anemia with 5-azacytidine results in increased fetal hemoglobin production and is associated with nonrandom hypomethylation of DNA around the γδβ-globin gene complex. Proc. Natl. Acad. Sci. USA **80:**4842–4846.
3. **Constantoulakis, P., B. Josephson, L. Mangahas, T. Papayannopoulou, T. Enver, F. Costantini, and G. Stamatoyannopoulos.** 1991. Locus control region-A gamma transgenic mice: a new model for studying the induction of fetal hemoglobin in the adult. Blood **77:**1326–1333.
4. **Costello, J., and P. Vertino.** 2002. Methylation matters: a new spin on maspin. Nat. Genet. **31:**123–124.
5. **Ehrlich, M., M. A. Gama-Sosa, L. H. Huang, R. M. Midgett, K. C. Kuo, R. A. McCune, and C. Gehrke.** 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. Nucleic Acids Res. **10:**2709–2721.
6. **Forsberg, E., K. Downs, H. Christensen, H. Im, P. A. Nuzzi, and E. H. Bresnick.** 2000. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. Proc. Natl. Acad. Sci. USA **97:**14494–14499.
7. **Futscher, B. W., M. M. Oshiro, R. J. Wozniak, N. Holtan, C. L. Hanigan, H. Duan, and F. E. Domann.** 2002. Role for DNA methylation in the control of cell type specific maspin expression. Nat. Genet. **31:**175–179.
8. **Hsiao, W. L., S. Gattoni-Celli, and I. B. Weinstein.** 1986. Effects of 5-azacytidine on expression of endogenous retrovirus-related sequences in C3H 10T1/2 cells. J. Virol. **57:**1119–1126.
9. **Hu, X., S. Eszterhas, N. Pallazzi, E. E. Bouhassira, J. Fields, O. Tanabe, S. A. Gerber, M. Bulger, J. D. Engel, M. Groudine, and S. Fiering.** 2006. Transcriptional interference among the murine β-like globin genes. Blood **109:**2210–2216.
10. **Hutchins, A. S., A. C. Mullen, H. W. Lee, K. J. Sykes, F. A. High, B. D. Hendrich, A. P. Bird, and S. L. Reiner.** 2002. Gene silencing quantitatively controls the function of a developmental trans-activator. Mol. Cell **10:**81–91.
11. **Jackson-Grusby, L., C. Beard, R. Possemato, M. Tudor, D. Fambrough, G. Csankovszki, J. Dausman, P. Lee, C. Wilson, E. Lander, and R. Jaenisch.** 2001. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat. Genet. **27:**31–39.
12. **Jaenisch, R., A. Schnieke, and K. Harbers.** 1985. Treatment of mice with 5-azacytodine efficiently activates silent retroviral genomes in different tissues. Proc. Natl. Acad. Sci. USA **82:**1451–1455.
13. **Jones, P. L., G. J. Veenstra, P. A. Wade, D. Vermaak, S. U. Kass, N. Landsberger, J. Strouboulis, and A. P. Wolffe.** 1998. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. **19:**187–191.
14. **Kina, T., K. Ikuta, E. Takayama, K. Wada, A. S. Majumdar, I. L. Weissman, and Y. Katsura.** 2000. The monoclonal antibody TER-119 recognizes a molecule associated with glycophorin A and specifically marks the late stages of murine erythroid lineage. Br. J. Haematol. **109:**280–287.
15. **Kingsley, P., J. Malik, R. Emerson, T. Bushnell, K. McGrath, L. Bloedorn, M. Bulger, and J. Palis.** 2006. "Maturational" globin switching in primary primitive erythroid cells. Blood **107:**1665–1672.
16. **Klose, R., and A. Bird.** 2006. Genomic DNA methylation: the mark and its mediators. Trends Biochem. Sci. **31:**89–97.
17. **Koshy, M., L. Dorn, L. Bressler, R. Molokie, D. Lavelle, N. Talischy, R. Hoffman, W. van Overveld, and J. DeSimone.** 2000. 2-Deoxy 5-azacytidine and fetal hemoglobin induction in sickle cell anemia. Blood **96:**2379–2384.
18. **Lewis, J. D., R. R. Meehan, W. J. Henzel, I. Maurer-Fogy, P. Jeppesen, F. Klein, and A. Bird.** 1992. Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. Cell **69:**905–914.
19. **Ley, T. J., J. DeSimone, N. P. Anagnou, G. H. Keller, R. K. Humphries, P. H. Turner, N. S. Young, P. Keller, and A. W. Nienhuis.** 1982. 5-Azacytidine selectively increases gamma-globin synthesis in a patient with β-thalassemia. N. Engl. J. Med. **301:**1469–1475.
20. **Lorincz, M. C., D. R. Dickerson, M. Schmitt, and M. Groudine.** 2004. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. Nat. Struct. Mol. Biol. **11:**1068–1075.
21. **Mavilio, F., A. Giampaolo, A. Care, G. Migliaccio, M. Calandrini, G. Russo, G. Pagliardi, G. Mastroberardino, M. Mainucci, and C. Peschle.** 1983. Mo-

lecular mechanisms of human hemoglobin switching: selective undermethylation and expression of globin genes in embryonic, fetal, and adult erythroblasts. Proc. Natl. Acad. Sci. USA **80:**6907–6911.
22. **McGrath, K., and J. Palis.** 2005. Hematopoisis in the yolk sac: more than meets the eye. Exp. Hematol. **33:**1021–1028.
23. **Meehan, R. R., J. D. Lewis, and A. P. Bird.** 1992. Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. Nucleic Acids Res. **20:**5085–5092.
24. **Nan, X., H. H. Ng, C. A. Johnson, C. D. Laherty, B. M. Turner, R. N. Eisenman, and A. Bird.** 1998. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature **393:**386–389.
25. **Oppenheim, A., Y. Katzir, E. Fibach, A. Goldfarb, and E. Rachmilewitz.** 1985. Hypomethylation of DNA derived from purified human erythoid cells correlates with gene activity of the β-globin cluster. Blood **66:**1202–1207.
26. **Pace, B., Q. Li, K. Peterson, and G. Stamatoyannopoulos.** 1994. α-Amino butyric acid cannot reactivate the silenced gamma gene of the beta locus YAC transgenic mouse. Blood **84:**4344–4353.
27. **Palis, J., and M. Yoder.** 2001. Yolk-sac hematopoiesis: the first blood cells of mouse and man. Exp. Hematol. **29:**927–936.
28. **Paulin, R., G. W. Griss, M. D. Davey, and A. A. Piper.** 1998. Urea improves efficiency of bisulfite mediated sequencing of 5′-methylcytosine in genomic DNA. Nucleic Acids Res. **26:**5009–5010.
29. **Rupon, J. W., S. Z. Wang, K. Gaensler, J. Lloyd, and G. D. Ginder.** 2006. Methyl binding domain protein 2 mediates gamma-globin gene silencing in adult human βYAC transgenic mice. Proc. Natl. Acad. Sci. USA **103:**6617–6622.
30. **Takai, D., and P. A. Jones.** 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. Proc. Natl. Acad. Sci. USA **99:**3740–3745.
31. **Tweedie, S., J. Charlton, V. Clark, and A. Bird.** 1997. Methylation of genomes and genes at the invertebrate-vertebrate boundary. Mol. Cell. Biol. **17:**1469–1475.
32. **van der Ploeg, L. H., and R. A. Flavell.** 1980. DNA methylation in the human gamma delta beta-globin locus in erythroid and nonerythroid tissues. Cell **19:**947–958.
33. **Whitelaw, E., P. Lamb, P. Hogben, and N. J. Proudfoot.** 1989. The globin switch at the level of mRNA in the developing mouse. Prog. Clin. Biol. Res. **316:**323–333.
34. **Whitney, J. B., III.** 1977. Differential control of the synthesis of two hemoglobin beta chains in normal mice. Cell **12:**863–871.
35. **Yoder, J. A., C. P. Walsh, and T. H. Bestor.** 1997. Cytosine methylation and the ecology of intragenomic parasites. Trends Genet. **13:**335–340.