Published in final edited form as:

Blood. 2002 November 1; 100(9): 3077–3086. doi:10.1182/blood-2002-04-1104.

Locus control regions

Qiliang Li, Kenneth R. Peterson, Xiangdong Fang, and George Stamatoyannopoulos
Division of Medical Genetics and Department of Genome Sciences, University of Washington, Seattle, and the Department of Biochemistry and Molecular Biology and the Department of Anatomy and Cell Biology, University of Kansas Medical Center, Kansas City

Abstract

Locus control regions (LCRs) are operationally defined by their ability to enhance the expression of linked genes to physiological levels in a tissue-specific and copy number-dependent manner at ectopic chromatin sites. Although their composition and locations relative to their cognate genes are different, LCRs have been described in a broad spectrum of mammalian gene systems, suggesting that they play an important role in the control of eukaryotic gene expression. The discovery of the LCR in the β -globin locus and the characterization of LCRs in other loci reinforces the concept that developmental and cell lineage-specific regulation of gene expression relies not on gene-proximal elements such as promoters, enhancers, and silencers exclusively, but also on long-range interactions of various *cis* regulatory elements and dynamic chromatin alterations.

Introduction

Locus control regions (LCRs) are operationally defined by their ability to enhance the expression of linked genes to physiological levels in a tissue-specific and copy number-dependent manner at ectopic chromatin sites. The components of an LCR commonly colocalize to sites of DNase I hypersensitivity (HS) in the chromatin of expressing cells. The core determinants at individual HSs are composed of arrays of multiple ubiquitous and lineage-specific transcription factor-binding sites.

The LCR was first identified in the human β -globin locus.¹ (For a review, see Stamatoyannopoulos and Grosfeld,² Fraser and Grosfeld,³ and Li et al.⁴) Early studies showed that a 5-kilobase (kb) β -globin gene segment, including a 1.5-kb promoter region, was expressed in erythroleukemia cell lines, implying that this fragment contains all the regulatory elements necessary for proper expression. However, this fragment did not uniformly promote gene expression in transgenic mice.^{5–7} The gene was expressed in only a small proportion of transgenic mice, but expression was far below physiologically significant levels and was variable between lines. These findings suggested that a major regulatory element required for reproducible, high-level expression *in vivo* was missing in this construct. Clues regarding the nature of the missing element came from several observations. For example, in some forms of β -thalassemia the genes of the β -globin locus are intact but not expressed.^{8,9} A defect common to the loci underlying these conditions was a large deletion upstream of the β -like globin genes. This deletion results in a closed chromatin conformation spanning the whole locus and leads to suppression of gene expression.^{8,10} Thus, these data suggested that the deleted DNA segment contained an indispensable *cis*-acting regulatory element required for β -globin expression *in vivo*. The existence of such a regulatory element was also implied by the presence of

developmentally stable, erythroid-specific HSs 6 to 20 kb 5' to the ϵ -globin gene.^{11,12} Definitive evidence for the presence of the LCR came from transgenic mouse studies.¹ Linkage of this region to a β -globin gene resulted in expression of the gene at a level comparable to the endogenous mouse β -globin genes in a position-independent, copy number-dependent manner. LCRs have been described in a broad spectrum of mammalian gene systems, suggesting that they play an important role in the control of eukaryotic gene expression.

Properties of LCRs

Transcriptional enhancer activity

The most prominent property of the LCRs is their strong, transcription-enhancing activity. The β -globin LCR is located 6 to 22 kb 5' to the first (embryonic) globin gene in the locus (Figure 1). It consists of 5 DNase I-hypersensitive sites, 5'HSs 1 to 5. HSs 1 to 4 are formed only in erythroid cells, while 5'HS5 is found in multiple lineages of cells, but it is not constitutive.¹³ When the LCR is absent, transcription of the human β -globin gene is usually less than 1% of the endogenous murine β -globin mRNA in transgenic mice, if it is expressed at all.⁵⁻⁷ Inclusion of the LCR increases β -globin gene expression to a level comparable to that of the mouse β -globin genes in all transgenic animals, indicating that the LCR has strong enhancer activity.¹ LCR enhancer activity is also significant at its endogenous location, as demonstrated by LCR-deletion experiments.¹⁴⁻¹⁶ These deletions in the native chromosomes of mouse or human cell lines severely reduce the expression of globin genes.

The enhancer activity of the β -globin LCR resides in 5'HS2, 3, and 4, but not in 5'HS1 or 5 (for a review, see Stamatoyannopoulos and Grosfeld,² Fraser and Grosfeld,³ Li et al.,⁴ and Hardison et al.¹⁷). 5'HS2 behaves as a classical enhancer; that is, its activity can be detected in transient transfection assays. Enhancer activity in 5'HS3 or 4 can be detected only when they are integrated into chromatin (for a review, see Hardison et al.¹⁷ and references therein). A requirement for chromosomal integration suggests that alteration of chromatin structure may be involved in propagating the enhancer activity of these 2 HSs. 5'HS5 functions as a chromatin insulator.¹⁸⁻²⁰ The function of 5'HS1 remains to be defined.

The enhancer activity of the β -globin LCR is tissue specific; that is, the expression of globin genes is confined to erythroid cells when linked to the β -globin LCR.^{1,21} In addition, the LCR is able to enhance expression of linked heterogeneous nonglobin gene promoters in erythrocytes. However, when a nonglobin gene was coupled to the LCR, ectopic expression was observed in some transgenic mice.²² In these instances, although the LCR conferred erythroid-specific gene expression on the heterogeneous gene, the natural function of the linked promoter allowed expression outside the erythroid compartment. Thus, tissue-specific control of basal transcription may reside in the promoter, as is the case for the globin genes, whereas tissue-specific enhancement of gene expression may be a property of the LCR. These data suggest that tissue specificity is not really an intrinsic property of the LCR but depends on both the LCR and the promoter that it interacts with.

Central to understanding the enhancer functions of β -globin LCR is the identification of the transcription factors mediating enhancer activity. Enhancer activity of 5'HSs 2-4 resides in a 200-bp to 300-bp core, which contains an array of binding sites for ubiquitous and erythroid-specific transactors. A conserved sequence within 5'HS2, TGCTGA(C/G)TCA(T/C), is critical for strong enhancer activity.^{23,24} This Maf recognition element (MARE) is bound by multiple homodimeric and heterodimeric transcription factors in vitro.²⁵ These factors include Maf homodimers, heterodimers containing a Maf subunit and another bZIP protein (NF-E2, Nrf1, Nrf2, Bach1, Bach2), and heterodimers lacking a Maf subunit (API).²⁶⁻³¹ NF-E2 is the major protein found in nuclear extracts from murine erythroleukemia (MEL) cells that binds the tandem MAREs of 5'HS2, and globin gene expression closely parallels the level of NF-E2

binding activity.³² The MEL cell line CB3, which lacks p45, is severely impaired in globin gene expression, and transcription can be rescued by expression of NF-E2.^{32,33} The erythroid-specific transactivator p45/NF-E2 binds directly and specifically to 5'HS2 in erythroleukemia cells and mouse fetal liver. Chromatin immunoprecipitation (ChIP) assay showed that specific recovery of the 5'HS2 sequences was dependent upon the presence of p45 and intact MARE sites within 5'HS2.³⁴ Investigation of the binding of the p45/p18 (MafK) heterodimer or other small Maf proteins within the globin locus showed that prior to induction of MEL cell differentiation, the LCR was occupied by small Maf proteins, and that during erythroid maturation, the NF-E2 complex was recruited to the LCR and the active globin promoters, even though the promoters do not contain MAREs. This differentiation-coupled recruitment of the NF-E2 complex correlates with a more than 100-fold increase in β -major globin transcription, but is not associated with a significant change in locus-wide histone H3 acetylation. Thus, the β -globin gene locus may exist in a constitutively open chromatin conformation before terminal differentiation, and the recruitment of the NF-E2 complex to the LCR and active promoters may be a rate-limiting step in the activation of β -globin gene expression.³⁵ While the *in vivo* association of NF-E2 and HS2 of the β LCR is confirmed by ChIP assay, a knockout of p45 gene does not inhibit globin gene expression.³⁶ The absence of phenotype in the p45 knockout mice is not due to the result of compensation by Nrf-2, a factor closely related to p45, as demonstrated by the study of a double knockout of p45 and Nrf-2, which also fails to interfere with expression of the α - and β -globin genes.³⁷ These observations suggest an interchangeable function between members of the cap'n collar (CNC) subfamily of bZIP transcription factors.

LCR functions may affect the basic transcription machinery directly. RNA polymerase II (pol II), one of the essential components of the eukaryotic transcription apparatus, was found to be associated with the β -globin LCR in a p45/NF-E2-independent manner, whereas its recruitment to the promoter required p45/NF-E2. These data suggest that pol II accesses the LCR and p45/NF-E2 induces long-range transfer of pol II to the promoter, resulting in transcriptional activation.³⁸

Copy number–dependent gene expression and chromatin domain-opening activity

Another property of the LCRs is their ability to confer position-independent, copy number–dependent expression on a linked gene.¹ Copy number–dependent expression is widely considered to be indicative of open chromatin structure, that is, DNA that is accessible to transcription factors. Involvement of the β -globin LCR in creating open chromatin was suggested from analysis of β -thalassemia mutants with deletion of the LCR.^{8,10} In the Hispanic form of β -thalassemia, an approximately 35-kb region upstream of 5'HS1 is deleted, but the remainder of the globin locus is intact. However, none of the globin genes is expressed. The deletion produces a closed chromatin conformation that spans the entire locus.¹⁰ Consistent with this, only the intact LCR (5'HS1–5) can provide position-independent chromatin-opening activity in single-copy transgenic mice carrying the entire β -globin locus.³⁹ When one of the HSs was deleted from the LCR, expression of the β -globin genes appeared to be sensitive to the position of integration. In transgenic mice carrying single copies of small, recombinant 5' HS-globin gene constructs, only 5'HS3 is able to confer copy number–dependent gene expression. This observation led to the conclusion that 5'HS3 possesses the dominant chromatin-opening activity of the β -globin LCR.⁴⁰ However, 5'HS3 chromatin-opening activity may not be dominant, since it appears to be dependent upon the constitution of the constructs.⁴¹ Formation of hypersensitivity is a result of interaction of multiple ubiquitous and erythroid-specific transacting factors in the HS regions.⁴²

Recent studies have established that the human CD2 LCR achieves position-independent expression in the T cells of transgenic mice by overcoming heterochromatin-mediated position

effect variegation (PEV).⁴³ Fluorescence in situ hybridization (FISH) was used to identify the sites of transgene integration in individual mouse lines and allowed a correlation between the type of position effects induced by such chromosomal locations and the DNA sequences required to overcome them. Transgenic mice carrying a CD2 minigene attached only to the 3' CD2 transcriptional enhancer (the CD2 HSs 1 and 2) exhibited variegated expression when the transgene integrated in the centromere. In contrast, mice carrying a transgene with additional 3' sequences (the CD2 HS3) showed no variegation even when the latter integrated in centromeric positions. This indicates that the CD2 HS3 functions in the establishment and/or maintenance of an open chromatin domain and that human CD2 LCR is able to overcome the gene repression imposed by constitutive centromeric heterochromatin.

In conclusion, the ability to confer copy number-dependent expression of a transgene is used to distinguish a DNA fragment functioning as an LCR rather than a transcriptional enhancer. This criterion has been employed in identification of all LCR or LCR-like elements.

Timing and origin of DNA replication

The mammalian genome is made up of defined zones that undergo DNA replication in a programmed manner during the S phase of the cell cycle. Studies of individual genes have demonstrated that there is a correlation between replication timing and gene expression.^{44,45} The human β -globin locus replicates late in most cell types, but replicates early in erythroid cells.^{46,47}

Data generated from transgenic mice by FISH analysis mapped and characterized the replication zone surrounding the human globin locus on chromosome 11. These results showed that the β -globin LCR region (5'HSs 1–5) was sufficient for directing replication timing in a developmentally specific manner in vivo.⁴⁸ The LCR (5'HSs 1–5) also plays a role in setting up regional erythroid-specific, open chromatin structure in transgenic mice, and this function is likely intertwined with the ability to direct early replication timing.⁴⁹ Although early replication is generally correlated with gene expression, it has not been possible to decipher the cause-and-effect relationship between these 2 parameters.⁴⁸ Other results using targeted deletion of the LCR (5'HSs 1–5) showed that early replication timing and an open chromatin structure do not, by themselves, guarantee high levels of globin transcription in erythroid cells.⁵⁰ Therefore, an as yet undefined class of *cis*-acting elements may play a role in mediating control of replication timing, independent of transcription.

Many studies have emphasized the relationship between early replication and globin transcription in erythroid cells. However, these replication elements within the LCR also function in nonex-pressing cell types. Thus, one of the major roles for replication timing control at the globin locus may be to set up late replication with its accompanying inactive chromatin structure in nonerythroid cells. In this manner, repression of background transcription may be achieved.⁴⁸ Perhaps this is accomplished by restricting the exposure of newly assembled nucleosomes to histone deacetylases, specifically during replication in late S phase.⁴⁹ Recent evidence supporting this hypothesis suggests that HDAC2 is preferentially associated with late replication foci.⁵¹ Further data are required to determine whether the effects on replication are general features of LCRs and whether these effects influence transcription or are secondary to it.

Histone modification and heterochromatin

Despite numerous studies on the role of the LCR in controlling β -globin gene expression, the mechanism of long-range transactivation by the LCRs is poorly understood. Several models (including looping, tracking, linking, topologic alterations, and modification of proteins associated with chromatin) have been invoked to explain the functions of LCR.^{4,52–55} All the

models, directly or indirectly, implicate the ability of LCRs to alter chromatin configuration and conformation.

The effects of LCRs on chromatin acetylation have been studied in different model systems. Function of the human growth hormone (hGH) LCR has been linked to specific patterns of core histone acetylation. The hGH locus consists of 5 genes expressed in either the pituitary or the placenta.⁵⁶ This LCR consists of 5 HSs: 2 pituitary-specific (HSI, HSII), 1 placenta-specific (HSIV), and 2 shared (HSIII, HSV). In the pituitary, the LCR is encompassed in a somatotrope-specific domain of hyperacetylated chromatin that extends from the most 5' LCR component to the hGH-N promoter. Further analysis shows that the hGH LCR, located 14.5 kb upstream from the hGH-N promoter, plays a critical, specific, and nonredundant role in facilitating promoter transacting factor binding and activation of hGH-N transcription. It also plays an essential role in establishing a 32-kb acetylated region that encompasses the entire hGH LCR contiguous with the hGH-N promoter. Separate positive elements in the LCR (HSI, HSII) for pituitary-expressed genes, or in gene-proximal sequences (P-elements) for placenta-expressed genes, activate their respective target genes by tissue-specific recruitment of different histone acetyltransferase activities, resulting in distinct patterns of acetylation across the locus.⁵⁷ These data support a model for long-range gene activation via LCR-mediated targeting and extensive spreading of core histone acetylation.⁵⁸

The functions of the LCR in the β -globin locus appear to be different from those of the hGH gene cluster. Deletion of 5'HS2–5 of the human β -globin LCR did not affect the general pattern of histone H4 acetylation of a β -globin locus transgene.⁵⁹ Other studies reported that although deletion of the murine β -globin LCR decreased the rate of β -globin transcription, it did not alter the acetylation status of histone H3 or H4 within the promoter region.⁶⁰ Thus, histone H3 or H4 acetylation at the β -globin promoter may be independent of LCR function.

NF-E2 is required for histone hyperacetylation at the adult β -globin promoter, but not at the LCR.^{38,61} Other data demonstrated that the β -globin LCR and transcriptionally active promoters were enriched in acetylated histones in fetal liver relative to fetal brain, whereas the inactive promoters were hypoacetylated. In contrast, the LCR and both active and inactive promoters were hyperacetylated in yolk sac. 5'HS2 was also hyperacetylated in murine ES cells, whereas β -globin promoters were hypoacetylated. Thus, the acetylation pattern varied at different developmental stages. Histone deacetylase inhibition selectively increased acetylation at a hypoacetylated promoter in fetal liver, suggesting that active deacetylation contributes to silencing of promoters. Therefore, dynamic histone acetylation and deacetylation activities may play an important role in the developmental control of β -globin gene expression.⁶¹

DNA methylation is important in mammalian development because it controls gene expression through chromatin closure and gene silencing. During development, gene loci expressed in a tissue-specific manner become selectively demethylated in the appropriate cell types by poorly understood processes. The LCRs may play a role in tissue-specific DNA demethylation. Studies of the methylation status of the LCR for the mouse T-cell receptor (TCR) α/δ locus support such a role. Tissue-specific functions of this LCR depend largely on 2 HSs, HS1 (T-cell receptor α enhancer) and HS1'. These HSs induce lymphoid organ-specific DNA demethylation in a region located 3.8 kb away, with little effect on intervening methylated DNA. Demethylation is impaired in mice with a germ line deletion of the HS1/HS1' clusters. Using 5'-deletion mutants of a transgenic LCR reporter gene construct, HS1' can act in the absence of HS1 to direct this tissue-specific DNA demethylation event. Therefore, elements of an LCR may control tissue-specific DNA methylation patterns both in transgenes and in native loci.⁶²

In vivo function of LCRs

As discussed above, LCRs possess all the properties necessary for opening a chromosome domain and preventing heterochromatinization at ectopic sites. This property of the LCR most prominently distinguishes it from enhancers. Thus, a broadly accepted model for the major role of the β -globin LCR in vivo is to open and/or maintain a permissive chromatin conformation within the β -globin locus in erythroid cells, although enhancement of transcription is also an essential function. Surprisingly, when the entire mouse β -globin LCR (5'HS1–6) was deleted by homologous recombination, the formation of the general DNase I-sensitivity associated with the β -globin locus domain was not affected; however, transcription of all β -like globin genes was strikingly reduced.^{14,63} These observations raise several questions regarding the real in vivo function of the β -globin LCR. Is this LCR simply another enhancer in the β -globin locus? If the β -globin LCR functions only as an enhancer within endogenous β -globin loci, how can this fact be reconciled with observations from transgenic mouse studies in which chromatin-opening activity is characteristic?

Understanding the in vivo function of the LCRs is associated with our knowledge of the process of gene activation. A prevailing model for gene activation is that it is a stepwise process. The first step is chromatin opening. Opening allows transacting transcription factors and cofactors to access chromatin and assemble a functional transcription apparatus. Genes in open chromatin domains are poised for expression. When protein activators are present, transcription commences and high-level gene expression is achieved. Chromatin opening is manifested by an increase in its sensitivity to DNase I or other nucleases. General DNase I sensitivity represents a level of sensitivity on the order of one magnitude greater than bulk chromatin. General sensitivity may stretch over regions of several hundred kilobases. Within these regions of general sensitivity, small regions (< 300 bp) of DNase I sensitivity may be 2 orders of magnitude more sensitive than bulk chromatin; these regions are termed DNase I-hypersensitive sites. Although increased DNase I sensitivity may be due to improved accessibility of DNA packed in chromatin, DNase I sensitivity indeed is an ambiguous indicative of chromatin structure. The precise molecular nature of the alterations underlying accessibility has not been delineated. Questions persist as to whether the changes occur at the 30-nm fiber level or at the nucleosome level and whether all histone tails in the general DNase I sensitive regions are modified (acetylated, methylated, or phosphorylated) in the same fashion. Although general DNase I sensitivity was detected in both normal and β -globin LCR knockout mice in erythroid cells, previous data do not indicate whether the general DNase I sensitivity detected in LCR knockout mice and that detected in normal mice represent identical or different chromatin configurations.

Regardless of what the chromatin configuration may be, chromatin of the globin locus is more sensitive than bulk DNA in LCR knockout mice. In the absence of the LCR, an alternate pathway for establishment and maintenance of open chromatin must exist. Chromatin is not a structurally inert entity. Most likely it undergoes many dynamic conformational transitions that may be important in facilitating interactions between transacting factors and DNA. DNA probably unwraps from the edge of the nucleosome, since sites within nucleosomal DNA are transiently separated from histones with a probability of 1 in 10^3 to 10^5 moving from the periphery of the nucleosome toward the center.^{64,65} Thus, given the dynamic nature of this system, factors present at sufficient concentrations and having high affinities for naked DNA may be able to compete efficiently with histone proteins for binding, thereby ensuring significant loading of these proteins at their cognate DNA elements in chromatin. Moreover, some transacting factors, such as GATA-4, are able to bind to compacted chromatin and open up a local chromatin.⁶⁶ Other transacting factors then attain an opportunity to access enhancer or promoter elements and further remodel chromatin by recruiting and targeting chromatin modifying and remodeling machinery. Since a large number of factor-binding sites are

scattered throughout the β -globin locus, particularly at promoters, they are able to recruit various proteins and cofactors in the erythroid environment. Accumulation of a large amount of small, qualitative changes may finally lead to a major change in chromatin structure. Such a synergistic mechanism could result in an open chromatin at low level in the absence of the LCR. Synergistic mechanisms have been postulated for transcription activation via cooperation of multiple transactivators⁶⁷ and for heterochromatin formation in a mass action model.⁶⁸ Based on this model, the LCR does not necessarily possess a specified chromatin opening activity.

The LCR chromatin-opening activity manifested in transgenic studies indeed results from the unique feature of the LCR that numerous binding sites clustered in the region induce an exponent (synergistic) effect on chromatin structure. Closed chromatin is considered the default status and is found in the vast majority of the chromosome. Thus, most transgenes are integrated in sites of closed chromatin. When transgene expression is detected, the chromatin region surrounding the transgene is invariably open. Although it is easy to surmise that chromatin opening is a prerequisite of gene transcription, this conclusion is not definitive. The 2 processes may be separated during in vitro assays, but more likely these 2 events are mutually interdependent in vivo. Thus, the chromatin-opening activity of the LCR manifested in transgenic studies must also function at the endogenous locus in vivo in a similar manner. Experimental description of chromatin opening indeed includes multiple distinctive states in chromatin conformation and configuration. Chromatin-opening activity has to be considered as an integrated but not necessarily linear event in gene activation.

Mechanisms of globin gene activation by the LCR

Clearly, the LCR has a role in enhancement of globin gene expression, although some uncertainty exists regarding the direct effects of the LCR on chromatin conformation. Analysis of LCR function at its endogenous location in cell lines suggests that it is limited to globin gene transcription activation,^{14,15} whereas transgenic experiments suggest that it is also necessary for the establishment and maintenance of an open β -globin chromatin domain. Regardless of whether the LCR functions in one or both of these processes, it does so over a long distance. Studies using both native loci and constructs in transgenic mice offer insights as to how the β -globin LCR accomplishes transcriptional activation. Four models of LCR function have been proposed: looping, tracking, facilitated tracking, and linking (Figure 2). Available data neither strongly support nor preclude any of them.

The looping model suggests that the 5'HSs of the β -globin LCR fold to form a holocomplex, with the HS core elements forming an active site that binds transcription factors and the core-flanking sequences constraining the holocomplex in the proper conformation (Figure 2A). This structure physically "loops," so that the LCR comes in close proximity to the appropriate promoter. Close association with gene-proximal promoter and enhancer elements allows the delivery of LCR-bound transcription proteins and other coactivators that interact with the basal transcription apparatus, already bound at the promoter to form a stable transcription complex, thus enhancing globin gene expression.^{39,52,69,70} A variation of this model suggests that the LCR initially serves as a multiple element receptor that acts as a hub for factor binding to direct chromatin remodeling.⁷¹ Once chromatin-remodeling activity has been completed, the LCR directly interacts with downstream genes to facilitate their expression.

Several data support the looping model. Deletion of the 5'HS2 core abolished expression of the ϵ -, γ -, and β -globin genes.⁷⁰ However, when the entire 5'HS2 region of conserved sequence similarity (core and flanking sequences) was removed, the ϵ -, γ - and β -globin genes were expressed in the correct temporal order, although the levels of each were decreased severalfold.⁷² These data suggested that, in the case of the core deletion, the 5'HS2 flanking regions were

able to interact with the flanking sequences of the remaining intact 5'HSs to form the normal holocomplex conformation. Removal of only the 5'HS2 core, in effect, destroyed the active site of the holocomplex, resulting in a dominant-negative mutation that crippled LCR function. In contrast, when the entire 5'HS2 region was deleted, the remaining 5'HS sites were able to adapt an alternate holocomplex conformation with a slightly less effective active site consisting of the remaining 5'HS cores and constrained in form by the remaining 5'HS flanking sequences.⁷² Similar results were found with 5'HS3 core deletions versus complete deletion of 5'HS3.^{72–74}

Further evidence supporting the notion of a holocomplex suggests that the LCR interacts with only 1 globin gene promoter at a time and that it may “flip-flop” between 2 or more promoters, depending on the stage of development.^{39,52} In this model, the LCR holocomplex is free to move from gene to gene. A parameter relevant to holocomplex function is the distance between the LCR and its target gene, which has been shown to affect the probability that these 2 elements will interact.^{52,75–77} This probability is constant for a gene at a specific stage of development. As development proceeds, the LCR has increasingly stable interactions with more distant globin genes, which is largely a function of the changing transcription factor milieu. Thus, it is mainly the availability of specific transcription factors and distance of a gene from the LCR that constrain the frequency of LCR-gene interactions during development.^{75–77}

In the tracking, or scanning, model, erythroid-specific and ubiquitous transcription factors and cofactors bind recognition sequences in the LCR sequences, forming an activation complex that migrates, or tracks, linearly along the DNA helix of the locus (Figure 2B).^{53,78} When this transcription complex encounters the basal transcription machinery located at the correct (according to the developmental stage) promoter, the complete transcriptional apparatus is assembled and transcription of that gene is initiated. If this model is valid, the expectation would be that some aberrant transcripts would arise from cryptic start sites along the locus. In fact, transcripts were detected across the LCR and intergenic regions in erythroid cells, but not in nonerythroid cells.^{53,79} However, these transcripts were nuclear-specific; they were not found in the cytoplasm, suggesting that they were not processed into mature messenger RNAs. The function of such intergenic transcription may be to deliver transcription complex proteins to the globin gene promoters via the tracking mechanism. Alternatively, it is possible that the function of these transcripts is to establish and maintain an open chromatin conformation permissive for gene transcription, although the persistence of DNase I sensitivity following deletion of the LCR in cell lines argues against this role.^{14,15} Deacetylases and methylases within the complex may reorganize chromatin after the transcription complex activates transcription, possibly to limit activation to a particular developmental stage.

The facilitated-tracking model incorporates aspects of both the looping and tracking models (Figure 2C).⁷⁸ An LCR bound-transcription factor and coactivator complex loops to contact downstream DNA in promoter-distal regions, where the transcription factor complex is released. This complex then tracks in small steps along the chromatin until it encounters the appropriate promoter with its associated bound proteins. A stable loop structure is established and gene expression proceeds.

The linking model proposes that chromatin facilitator proteins bound throughout the locus define the domain to be transcribed and mediate the sequential stage-specific binding of transcription factors (Figure 2D).⁵⁵ Non-DNA-binding facilitator proteins form a continuous protein chain from the LCR to the globin gene to be transcribed, linking proteins bound at a transcriptionally primed gene to one another.⁵⁴ Support for this model comes from the *Drosophila* Chip protein complex.⁸⁰ Chip protein complexes interact with transcription factors bound at a promoter region at a specific developmental time point. The Chip-tagged promoter is targeted for transcriptional activation. It was speculated that a homologous mammalian

protein complex may act as the facilitating guide for transcription initiation, associating with transcription factors in the globin gene promoter regions and aiding gene activation.⁵⁵ This Chip-like protein complex may allow transcriptional activation of one gene at a time, while simultaneously blocking transcription outside of the region, accounting for the developmental stage-specific expression of the β -like globin genes. The β -globin locus may have several transcription factor-bound promoter regions linked in a chainlike fashion. Chip-like proteins then dissociate and move to another promoter link to target that promoter for LCR interaction. Thus, globin gene switching proceeds.

LCRs in other systems

Several elements have been characterized in mammals (including humans, mice, rats, chickens, rabbits, sheep, and goats) that meet the criteria for LCR function (Table 1). Several other elements have been identified that likely will be confirmed as LCRs, including one in the medaka fish tyrosinase gene.^{81–84} Structurally, these LCRs are composed of varying numbers of tissue-specific DNase I-hypersensitive sites. The HSs of nonglobin LCRs have been extensively characterized and consist of a 150- to 300-bp central core containing a high density of transcription factor binding sites.^{85–87} Although the β -globin LCR consists of 5 HSs clustered on one contiguous piece of DNA, the sequences that embody a complete LCR do not have to be located together, whether upstream of, downstream of, or within the genes they control. Other LCRs are a collection of elements with different numbers of HSs spread over large distances. The relative simplicity of the β -globin LCR with regard to its single group of HSs may have contributed to its early discovery. Identification of LCRs in complex multigene loci, where the elements are interspersed among the genes, is a difficult task. Functionally, they all exhibit some or all of the properties associated with the β -globin LCR, most commonly the hallmark of copy number-dependent, site-of-integration-independent expression of their cognate loci or linked transgenes.

Most of the data regarding LCR function have come from studies of the human and murine β -globin LCRs. Several insights have been gained with studies of the chicken LCR. Organization of the chicken LCR is similar to its human counterpart except that one of the LCR elements is located between the adult β^A and embryonic ϵ genes (the β^A/ϵ enhancer).⁸⁸ This enhancer is able to confer site-independent expression to the chicken β^A -globin gene in transgenic mice.⁸⁹ Chromatin unfolding of the chicken β -locus requires the presence of both the LCR and the promoter.⁹⁰ Chicken HS4 demarcates the 5' border of the locus, which functions as a powerful chromatin insulator.⁹¹ The insulating function of chicken HS4 is manifested by enhancer blocking activity and position-effect protection. These two activities are separable:⁹² the former is mediated by a transacting factor CTCF⁹³ and the latter function may be achieved by highly efficient recruitment of histone acetyltransferase by the HS4 element.⁹⁴

Some novel data regarding LCR function have come from studies of nonglobin LCRs. As previously discussed, the human CD2 LCR was shown to be essential for establishing an open chromatin configuration, even in the absence of enhancer activity.^{43,95} Thus, LCRs appear to operate by ensuring an open chromatin configuration. As discussed earlier, the T cell-specific TCR α/δ (TCR α) LCR has been implicated in tissue-specific DNA demethylation, an important role for LCRs, since DNA methylation may cause chromatin closure and gene silencing. Additional information regarding LCR function was obtained from studies of this LCR. The TCR α LCR consists of 8 HSs located downstream of the T-cell receptor (TCR) gene.⁹⁶ It is a bifunctional element, regulating both the TCR gene and the adjacent, ubiquitously expressed *Dad1* antiapoptosis gene. Two subregions of the TCR α LCR were identified: one that constituted a novel non-tissue-restricted chromatin-opening element and one an immediate upstream sequence comprising the 4 proximal HSs that restored tissue specificity to the

downstream chromatin-opening element.⁹⁷ The HSs of this tissue-specificity region map near 2 transcriptional silencers, the TCR α enhancer (HS1) and a region of unknown function; the region between the enhancer and the unknown HS appears to be responsible for the tissue specificity.⁹⁸ The proximal tissue-specific element may insulate the TCR gene from the LCR in other tissues, without affecting the TCR α LCR-*Dad1* interaction. The occurrence of activators and insulators in LCRs appears to be a common theme, suggesting that the interaction of these elements may modulate LCR function. In fact, the tissue-unrestricted HS element suppressed PEV of a linked transgene in a wide variety of tissues and was bound by several ubiquitously expressed transcription factors.⁹⁹ However, when the full-length LCR was present, tissue-specific binding of tissue-restricted proteins was observed, demonstrating that a widely active LCR element can interact synergistically with other LCR elements to produce tissue-specific LCR activity via differential protein binding.

Other results suggest that LCRs may, in some instances, activate gene expression through a mechanism that includes increased histone acetylation. A cassette derived from 4 HSs of the 3' murine immunoglobulin heavy chain (IgH) locus LCR was linked to the *c-myc* gene.¹⁰⁰ This LCR mediated a widespread increase in acetylation, not only within the promoter region of the *c-myc* gene, but also over substantial distances upstream and downstream of the transcription site. Studies of the hGH LCR described earlier suggest that LCRs may increase histone acetylation by targeted recruitment and subsequent spreading of histone acetyltransferase activity to encompass and activate remote target genes.¹⁰¹

Two elements of the T-cell receptor γ (TCR γ) locus, a 3' enhancer (3'E (C γ 1)) and a region called HsA located between the V γ 5 and V γ 2 genes, constitute an LCR.¹⁰² HsA alone supported position-independent transcription in mature, but not immature, T cells; 3'E (C γ 1) alone supported position-dependent expression in both immature and mature cells. Copy number-dependent, position-independent expression was obtained at both stages of development, suggesting that HsA provides chromatin-opening activity. In addition, HsA was required for rearrangement of transgenic recombination substrates, an essential component of TCR loci.

Another interesting variant of the LCR theme is the human keratin-18 (K18) gene, which contains a 323-bp fragment that confers position-independent, copy number-dependent expression upon heterologous transgenes.¹⁰³ This fragment is composed primarily of an Alu repetitive element that was partly responsible for the protective effects of sequences flanking K18, perhaps through its pol III transcriptional potential and inhibition of transcriptional interference from neighboring genes. Thus, Alu elements may function as regulatory domains within LCRs.

A pair of DNase I hypersensitive sites (site II) of the murine γ 1 heavy-chain gene lie approximately 2 kb 3' of the γ 1 promoter and exon 1 and just 5' of the γ 1 switch region. Site II functions as an LCR, conferring insertion site independence and copy number dependence on linked transgene expression, the hallmark of LCRs.¹⁰⁴ Messenger RNA is induced from γ 1 transgenes lacking site II by interleukin 4 (IL-4) and by CD40 ligation (CD40 ligand-CD40 interaction). However, in the absence of site II, the induction of transgenic RNA expression by CD40 ligation was greater than expected, suggesting that the elements within site II also participate in negative regulation of the number of germ line transcripts after CD40 ligation, an effect opposite to the enhancement of transcription observed with most LCRs.

Conclusions

LCRs have been identified in various loci of vertebrates. While their composition and location relative to their cognate genes are different, they share the common property of maintaining

physiological levels of gene expression, either at their natural position or at ectopic sites. This feature highlights the complexity of gene regulation. Developmental and cell lineage-specific regulation of gene expression relies not upon gene-proximal elements such as promoters, enhancers, and silencers exclusively, but also upon long-range interactions of various *cis* regulatory elements and dynamic chromatin alterations. The discovery of the LCR in the β -globin locus and the characterization of LCRs in other loci reinforces the need to study *in vivo* transcriptional regulation in the context of whole loci, so that essential regulatory elements are not excluded or overlooked.

Acknowledgments

Supported by National Institutes of Health grants DK53510, HL67336, HL20899, DK61805, and DK61804 and a Faculty Scholar Award from the Madison and Lila Self Graduate Fellowship awarded to K.R.P.

References

1. Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 1987;51:975–985. [PubMed: 3690667]
2. Stamatoyannopoulos, G.; Grosveld, F. Hemoglobin switching. In: Stamatoyannopoulos, G.; Majerus, PW.; Perlmutter, RM.; Varmus, H., editors. *Molecular Basis of Blood Diseases*. 3rd ed.. WB Saunders; Philadelphia, PA: 2001. p. 135–182.
3. Fraser P, Grosveld F. Locus control regions, chromatin activation and transcription. *Curr Opin Cell Biol* 1998;10:361–365. [PubMed: 9640537]
4. Li Q, Harju S, Peterson KR. Locus control regions: coming of age at a decade plus. *Trends Genet* 1999;15:403–408. [PubMed: 10498936]
5. Magram J, Chada K, Costantini F. Developmental regulation of a cloned adult beta-globin gene in transgenic mice. *Nature* 1985;315:338–340. [PubMed: 3858676]
6. Townes TM, Lingrel JB, Chen HY, Brinster RL, Palmiter RD. Erythroid-specific expression of human beta-globin genes in transgenic mice. *EMBO J* 1985;4:1715–1723. [PubMed: 2992937]
7. Kollias G, Wrighton N, Hurst J, Grosveld F. Regulated expression of human A gamma-, beta-, and hybrid gamma beta-globin genes in transgenic mice: manipulation of the developmental expression patterns. *Cell* 1986;46:89–94. [PubMed: 3719696]
8. Kioussis D, Vanin E, deLange T, Flavell RA, Grosveld FG. Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. *Nature* 1983;306:662–666. [PubMed: 6318113]
9. Driscoll MC, Dobkin CS, Alter BP. Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. *Proc Natl Acad Sci U S A* 1989;86:7470–7474. [PubMed: 2798417]
10. Forrester WC, Epner E, Driscoll MC, et al. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. *Genes Dev* 1990;4:1637–1649. [PubMed: 2249769]
11. Tuan D, Solomon W, Li Q, London IM. The “beta-like-globin” gene domain in human erythroid cells. *Proc Natl Acad Sci U S A* 1985;82:6384–6388. [PubMed: 3879975]
12. Forrester WC, Thompson C, Elder JT, Groudine M. A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc Natl Acad Sci U S A* 1986;83:1359–1363. [PubMed: 3456593]
13. Li Q, Zhang M, Duan Z, Stamatoyannopoulos G. Structural analysis and mapping of DNase I hypersensitivity of HS5 of the beta-globin locus control region. *Genomics* 1999;61:183–193. [PubMed: 10534403]
14. Epner E, Reik A, Cimborra D, et al. The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. *Mol Cell* 1998;2:447–455. [PubMed: 9809066]
15. Reik A, Telling A, Zitnik G, Cimborra D, Epner E, Groudine M. The locus control region is necessary for gene expression in the human beta-globin locus but not the maintenance of an open chromatin structure in erythroid cells. *Mol Cell Biol* 1998;18:5992–6000. [PubMed: 9742116]

16. Bender MA, Bulger M, Close J, Groudine M. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. *Mol Cell* 2000;5:387–393. [PubMed: 10882079]
17. Hardison R, Slightom JL, Gumucio DL, Goodman M, Stojanovic N, Miller W. Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. *Gene* 1997;205:73–94. [PubMed: 9461381]
18. Li Q, Stamatoyannopoulos G. Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. *Blood* 1994;84:1399–1401. [PubMed: 8068937]
19. Li Q, Zhang M, Han H, Rohde A, Stamatoyannopoulos G. Evidence that DNase I hypersensitive site 5 of the human beta-globin LCR functions as a chromosomal insulator in transgenic mice. *Nucleic Acids Res* 2002;30:2484–2491. [PubMed: 12034837]
20. Farrell CM, West AG, Felsenfeld G. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol Cell Biol* 2002;22:3820–3831. [PubMed: 11997516]
21. Blom van Assendelft G, Hanscombe O, Grosveld F, Greaves DR. The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 1989;56:969–977. [PubMed: 2924354]
22. Tewari R, Gillemans N, Harper A, et al. The human beta-globin locus control region confers an early embryonic erythroid-specific expression pattern to a basic promoter driving the bacterial lacZ gene. *Development* 1996;122:3991–3999. [PubMed: 9012519]
23. Ney PA, Sorrentino BP, McDonagh KT, Nienhuis AW. Tandem AP-1-binding sites within the human beta-globin dominant control region function as an inducible enhancer in erythroid cells. *Genes Dev* 1990;4:993–1006. [PubMed: 2116990]
24. Talbot D, Grosveld F. The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. *EMBO J* 1991;10:1391–1398. [PubMed: 1902783]
25. Motohashi H, Shavit JA, Igarashi K, Yamamoto M, Engel JD. The world according to Maf. *Nucleic Acids Res* 1997;25:2953–2959. [PubMed: 9224592]
26. Mignotte V, Wall L, deBoer E, Grosveld F, Romeo PH. Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res* 1989;17:37–54. [PubMed: 2911469]
27. Andrews NC, Erdjument-Bromage H, Davidson MB, Tempst P, Orkin SH. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 1993;362:722–728. [PubMed: 8469283]
28. Caterina JJ, Donze D, Sun CW, Ciavatta DJ, Townes TM. Cloning and functional characterization of LCR-F1: a bZIP transcription factor that activates erythroid-specific, human globin gene expression. *Nucleic Acids Res* 1994;22:2383–2391. [PubMed: 8036168]
29. Chan JY, Han XL, Kan YW. Cloning of Nrf1, an NF-E2-related transcription factor, by genetic selection in yeast. *Proc Natl Acad Sci U S A* 1993;90:11371–11375. [PubMed: 8248256]
30. Moi P, Chan K, Asunis I, Cao A, Kan YW. Isolation of NF-E2-related factor 2 (Nrf2), a NF-E2-like basic leucine zipper transcriptional activator that binds to the tandem NF-E2/AP1 repeat of the beta-globin locus control region. *Proc Natl Acad Sci U S A* 1994;91:9926–9930. [PubMed: 7937919]
31. Oyake T, Itoh K, Motohashi H, et al. Bach proteins belong to a novel family of BTB-basic leucine zipper transcription factors that interact with MafK and regulate transcription through the NF-E2 site. *Mol Cell Biol* 1996;16:6083–6095. [PubMed: 8887638]
32. Kotkow KJ, Orkin SH. Dependence of globin gene expression in mouse erythroleukemia cells on the NF-E2 heterodimer. *Mol Cell Biol* 1995;15:4640–4647. [PubMed: 7623856]
33. Lu SJ, Rowan S, Bani MR, Ben-David Y. Retroviral integration within the Fli-2 locus results in inactivation of the erythroid transcription factor NF-E2 in Friend erythroleukemias: evidence that NF-E2 is essential for globin expression. *Proc Natl Acad Sci U S A* 1994;91:8398–8402. [PubMed: 8078893]
34. Forsberg EC, Downs KM, Bresnick EH. Direct interaction of NF-E2 with hypersensitive site 2 of the beta-globin locus control region in living cells. *Blood* 2000;96:334–339. [PubMed: 10891470]

35. Sawado T, Igarashi K, Groudine M. Activation of beta-major globin gene transcription is associated with recruitment of NF-E2 to the beta-globin LCR and gene promoter. *Proc Natl Acad Sci U S A* 2001;98:10226–10231. [PubMed: 11517325]
36. Shivdasani RA, Orkin SH. Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2. *Proc Natl Acad Sci U S A* 1995;92:8690–8694. [PubMed: 7567998]
37. Martin F, van Deursen JM, Shivdasani RA, Jackson CW, Troutman AG, Ney PA. Erythroid maturation and globin gene expression in mice with combined deficiency of NF-E2 and nrf-2. *Blood* 1998;91:3459–3466. [PubMed: 9558405]
38. Johnson KD, Christensen HM, Zhao B, Bresnick EH. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol Cell* 2001;8:465–471. [PubMed: 11545748]
39. Milot E, Strouboulis J, Trimborn T, et al. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 1996;87:105–114. [PubMed: 8858153]
40. Ellis J, Tan-Un KC, Harper A, et al. A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human beta-globin locus control region. *EMBO J* 1996;15:562–568. [PubMed: 8599939]
41. Li Q, Stamatoyannopoulos JA. Position independence and proper developmental control of gamma-globin gene expression require both a 5' locus control region and a downstream sequence element. *Mol Cell Biol* 1994;14:6087–6096. [PubMed: 8065342]
42. Goodwin AJ, McInerney JM, Glander MA, Pomerantz O, Lowrey CH. In vivo formation of a human beta-globin locus control region core element requires binding sites for multiple factors including GATA-1, NF-E2, erythroid Kruppel-like factor, and Sp1. *J Biol Chem* 2001;276:26883–26892. [PubMed: 11304527]
43. Festenstein R, Tolaini M, Corbella P, et al. Locus control region function and heterochromatin-induced position effect variegation. *Science* 1996;271:1123–1125. [PubMed: 8599090]
44. Holmquist GP. Role of replication time in the control of tissue-specific gene expression. *Am J Hum Genet* 1987;40:151–173. [PubMed: 3551593]
45. Selig S, Okumura K, Ward DC, Cedar H. Delineation of DNA replication time zones by fluorescence in situ hybridization. *EMBO J* 1992;11:1217–1225. [PubMed: 1547781]
46. Epner E, Forrester WC, Groudine M. Asynchronous DNA replication within the human beta-globin gene locus. *Proc Natl Acad Sci U S A* 1988;85:8081–8085. [PubMed: 3186709]
47. Dhar V, Skoultschi AI, Schildkraut CL. Activation and repression of a beta-globin gene in cell hybrids is accompanied by a shift in its temporal replication. *Mol Cell Biol* 1989;9:3524–3532. [PubMed: 2796994]
48. Simon I, Tenzen T, Mostoslavsky R, Fibach E, Lande L, Milot E, Gribnau J, Grosveld F, Fraser P, Cedar H. Developmental regulation of DNA replication timing at the human beta globin locus. *EMBO J* 2001;20:6150–6157. [PubMed: 11689454]
49. Allshire R, Bickmore W. Pausing for thought on the boundaries of imprinting. *Cell* 2000;102:705–708. [PubMed: 11030613]
50. Simon, I.; Cedar, H. Temporal order of DNA replication. In: De Pamphilis, ML., editor. *DNA replication in eukaryotic cells*. Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY: 1996. p. 387–408.
51. Rountree MR, Bachman KE, Baylin SB. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 2000;25:269–277. [PubMed: 10888872]
52. Wijgerde M, Grosveld F, Fraser P. Transcription complex stability and chromatin dynamics in vivo. *Nature* 1995;377:209–213. [PubMed: 7675106]
53. Tuan D, Kong S, Hu K. Transcription of the hypersensitive Site HS2 enhancer in erythroid cells. *Proc Natl Acad Sci U S A* 1992;89:11219–11223. [PubMed: 1454801]
54. Engel JD, Tanimoto K. Looping, linking, and chromatin activity; new insights into beta-globin locus regulation. *Cell* 2000;100:499–502. [PubMed: 10721987]
55. Bulger M, Groudine M. Looping versus linking: toward a model for long-distance gene activation. *Genes Dev* 1999;13:2465–2477. [PubMed: 10521391]
56. Su Y, Liebhaber SA, Cooke NE. The human growth hormone gene cluster locus control region supports position-independent pituitary- and placenta-specific expression in the transgenic mouse. *J Biol Chem* 2000;275:7902–7909. [PubMed: 10713106]

57. Elefant F, Su Y, Liebhaber SA, Cooke NE. Patterns of histone acetylation suggest dual pathways for gene activation by a bifunctional locus control region. *EMBO J* 2000;19:6814–6822. [PubMed: 11118216]
58. Ho Y, Elefant F, Cooke N, Liebhaber S. A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. *Mol Cell* 2002;9:291–302. [PubMed: 11864603]
59. Schubeler D, Francastel C, Cimborra DM, Reik A, Martin DI, Groudine M. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev* 2000;14:940–950. [PubMed: 10783166]
60. Schubeler D, Groudine M, Bender MA. The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. *Proc Natl Acad Sci U S A* 2001;98:11432–11437. [PubMed: 11553791]
61. Forsberg EC, Downs KM, Christensen HM, Im H, Nuzzi PA, Bresnick EH. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc Natl Acad Sci U S A* 2000;97:14494–14499. [PubMed: 11121052]
62. Santoso B, Ortiz BD, Winoto A. Control of organ-specific demethylation by an element of the T-cell receptor-alpha locus control region. *J Biol Chem* 2000;275:1952–1958. [PubMed: 10636897]
63. Bender MA, Mehaffey MG, Telling A, et al. Independent formation of DNaseI hypersensitive sites in the murine beta-globin locus control region. *Blood* 2000;95:3600–3604. [PubMed: 10828050]
64. Anderson JD, Widom J. Sequence and position-dependence of the equilibrium accessibility of nucleosomal DNA target sites. *J Mol Biol* 2000;296:979–987. [PubMed: 10686097]
65. Polach KJ, Lowary PT, Widom J. Effects of core histone tail domains on the equilibrium constants for dynamic DNA site accessibility in nucleosomes. *J Mol Biol* 2000;298:211–223. [PubMed: 10764592]
66. Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol Cell* 2002;9:279–289. [PubMed: 11864602]
67. Carey M. The enhanceosome and transcriptional synergy. *Cell* 1998;92:5–8. [PubMed: 9489694]
68. Locke J, Kotarski MA, Tartof KD. Dosage-dependent modifiers of position effect variegation in *Drosophila* and a mass action model that explains their effect. *Genetics* 1988;120:181–198. [PubMed: 3146523]
69. Gribnau J, de Boer E, Trimborn T, et al. Chromatin interaction mechanism of transcriptional control in vivo. *EMBO J* 1998;17:6020–6027. [PubMed: 9774345]
70. Bungert J, Tanimoto K, Patel S, Liu Q, Fear M, Engel JD. Hypersensitive site 2 specifies a unique function within the human beta-globin locus control region to stimulate globin gene transcription. *Mol Cell Biol* 1999;19:3062–3072. [PubMed: 10082573]
71. Grosveld F. Activation by locus control regions? *Curr Opin Genet Dev* 1999;9:152–157. [PubMed: 10322132]
72. Peterson KR, Clegg CH, Navas PA, Norton EJ, Kimbrough TG, Stamatoyannopoulos G. Effect of deletion of 5'HS3 or 5'HS2 of the human beta-globin locus control region on the developmental regulation of globin gene expression in beta-globin locus yeast artificial chromosome transgenic mice. *Proc Natl Acad Sci U S A* 1996;93:6605–6609. [PubMed: 8692864]
73. Navas PA, Peterson KR, Li Q, et al. Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. *Mol Cell Biol* 1998;18:4188–4196. [PubMed: 9632803]
74. Bungert J, Dave U, Lim KC, et al. Synergistic regulation of human beta-globin gene switching by locus control region elements HS3 and HS4. *Genes Dev* 1995;9:3083–3096. [PubMed: 8543153]
75. Hanscombe O, Whyatt D, Fraser P, et al. Importance of globin gene order for correct developmental expression. *Genes Dev* 1991;5:1387–1394. [PubMed: 1714415]
76. Peterson KR, Stamatoyannopoulos G. Role of gene order in developmental control of human gamma- and beta-globin gene expression. *Mol Cell Biol* 1993;13:4836–4843. [PubMed: 8336720]
77. Dillon N, Trimborn T, Strouboulis J, Fraser P, Grosveld F. The effect of distance on long-range chromatin interactions. *Mol Cell* 1997;1:131–139. [PubMed: 9659910]
78. Blackwood EM, Kadonaga JT. Going the distance: a current view of enhancer action. *Science* 1998;281:61–63.

79. Ashe HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ. Intergenic transcription and transinduction of the human beta-globin locus. *Genes Dev* 1997;11:2494–2509. [PubMed: 9334315]
80. Dorsett D. Distant liaisons: long-range enhancer-promoter interactions in *Drosophila*. *Curr Opin Genet Dev* 1999;9:505–514. [PubMed: 10508687]
81. Fu L, Mambrini M, Perrot E, Chourrout D. Stable and full rescue of the pigmentation in a medaka albino mutant by transfer of a 17 kb genomic clone containing the medaka tyrosinase gene. *Gene* 2000;241:205–211. [PubMed: 10675031]
82. Mathieu N, Hempel WM, Spicuglia S, Verthuy C, Ferrier P. Chromatin remodeling by the T cell receptor (TCR)-beta gene enhancer during early T cell development: implications for the control of TCR-beta locus recombination. *J Exp Med* 2000;192:625–636. [PubMed: 10974029]
83. Agarwal S, Rao A. Long-range transcriptional regulation of cytokine gene expression. *Curr Opin Immunol* 1998;10:345–352. [PubMed: 9638372]
84. Yui MA, Hernandez-Hoyos G, Rothenberg EV. A new regulatory region of the IL-2 locus that confers position-independent transgene expression. *J Immunol* 2001;166:1730–1739. [PubMed: 11160218]
85. Bonifer C, Vidal M, Grosveld F, Sippel AE. Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. *EMBO J* 1990;9:2843–2848. [PubMed: 2390972]
86. Aronow BJ, Silbiger RN, Dusing MR, et al. Functional analysis of the human adenosine deaminase gene thymic regulatory region and its ability to generate position-independent transgene expression. *Mol Cell Biol* 1992;12:4170–4185. [PubMed: 1508212]
87. Kushida MM, Dey A, Zhang XL, et al. A 150-base pair 5' region of the MHC class I HLA-B7 gene is sufficient to direct tissue-specific expression and locus control region activity: the alpha site determines efficient expression and in vivo occupancy at multiple cis-active sites throughout this region. *J Immunol* 1997;159:4913–4929. [PubMed: 9366417]
88. Felsenfeld G. Chromatin structure and the expression of globin-encoding genes. *Gene* 1993;135:119–124. [PubMed: 8276248]
89. Reitman M, Lee E, Westphal H, Felsenfeld G. Site-independent expression of the chicken beta A-globin gene in transgenic mice. *Nature* 1990;348:749–752. [PubMed: 2175398]
90. Reitman M, Lee E, Westphal H, Felsenfeld G. An enhancer/locus control region is not sufficient to open chromatin. *Mol Cell Biol* 1993;13:3990–3998. [PubMed: 8321206]
91. Chung JH, Whiteley M, Felsenfeld G. A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* 1993;74:505–514. [PubMed: 8348617]
92. Recillas-Targa F, Pikaart MJ, Burgess-Beusse B, et al. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A* 2002;99:6883–6888. [PubMed: 12011446]
93. Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 1999;98:387–396. [PubMed: 10458613]
94. Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 2001;293:2453–2455. [PubMed: 11498546]
95. Greaves DR, Wilson FD, Lang G, Kioussis D. Human CD2 3'-flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. *Cell* 1989;56:979–986. [PubMed: 2564317]
96. Diaz P, Cado D, Winoto A. A locus control region in the T cell receptor alpha/delta locus. *Immunity* 1994;1:207–217. [PubMed: 7889409]
97. Ortiz BD, Cado D, Chen V, Diaz PW, Winoto A. Adjacent DNA elements dominantly restrict the ubiquitous activity of a novel chromatin-opening region to specific tissues. *EMBO J* 1997;16:5037–5045. [PubMed: 9305645]
98. Ortiz BD, Cado D, Winoto A. A new element within the T-cell receptor alpha locus required for tissue-specific locus control region activity. *Mol Cell Biol* 1999;19:1901–1909. [PubMed: 10022877]
99. Ortiz BD, Harrow F, Cado D, Santoso B, Winoto A. Function and factor interactions of a locus control region element in the mouse T cell receptor-alpha/Dad1 gene locus. *J Immunol* 2001;167:3836–3845. [PubMed: 11564801]

100. Madisen L, Krumm A, Hebbes TR, Groudine M. The immunoglobulin heavy chain locus control region increases histone acetylation along linked c-myc genes. *Mol Cell Biol* 1998;18:6281–6292. [PubMed: 9774645]
101. Elefant F, Cooke NE, Liebhaber SA. Targeted recruitment of histone acetyltransferase activity to a locus control region. *J Biol Chem* 2000;275:13627–13834.
102. Baker JE, Kang J, Xiong N, Chen T, Cado D, Raulet DH. A novel element upstream of the Vgamma2 gene in the murine T cell receptor gamma locus cooperates with the 3' enhancer to act as a locus control region. *J Exp Med* 1999;190:669–679. [PubMed: 10477551]
103. Willoughby DA, Vilalta A, Oshima RG. An Alu element from the K18 gene confers position-independent expression in transgenic mice. *J Biol Chem* 2000;275:759–766. [PubMed: 10625605]
104. Adams K, Ackerly H, Cunningham K, Dunnick W. A DNase I hypersensitive site near the murine gamma1 switch region contributes to insertion site independence of transgenes and modulates the amount of transcripts induced by CD40 ligation. *Int Immunol* 2000;12:1705–1713. [PubMed: 11099310]
105. Dang Q, Walker D, Taylor S, et al. Structure of the hepatic control region of the human apolipoprotein E/C-I gene locus. *J Biol Chem* 1995;270:22577–22585. [PubMed: 7673250]
106. Friend WC, Clapoff S, Landry C, et al. Cell-specific expression of high levels of human S100 beta in transgenic mouse brain is dependent on gene dosage. *J Neurosci* 1992;12:4337–4346. [PubMed: 1432098]
107. Jones BK, Monks BR, Liebhaber SA, Cooke NE. The human growth hormone gene is regulated by a multicomponent locus control region. *Mol Cell Biol* 1995;15:7010–7021. [PubMed: 8524268]
108. Kaios M, Fournier RE. Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. *Mol Cell Biol* 1995;15:198–207. [PubMed: 7799927]
109. Knotts S, Rindt H, Robbins J. Position independent expression and developmental regulation is directed by the beta myosin heavy chain gene's 5' upstream region in transgenic mice. *Nucleic Acids Res* 1995;23:3301–3309. [PubMed: 7667107]
110. Madisen L, Groudine M. Identification of a locus control region in the immunoglobulin heavy-chain locus that deregulates c-myc expression in plasmacytoma and Burkitt's lymphoma cells. *Genes Dev* 1994;8:2212–2226. [PubMed: 7958890]
111. Mills FC, Harindranath N, Mitchell M, Max EE. Enhancer complexes located downstream of both human immunoglobulin Calpha genes. *J Exp Med* 1997;186:845–858. [PubMed: 9294139]
112. Neznanov N, Thorey IS, Cecena G, Oshima RG. Transcriptional insulation of the human keratin 18 gene in transgenic mice. *Mol Cell Biol* 1993;13:2214–2223. [PubMed: 7681143]
113. Schmidt CM, Ehlenfeldt RG, Athanasiou MC, et al. Extraembryonic expression of the human MHC class I gene HLA-G in transgenic mice. Evidence for a positive regulatory region located 1 kilobase 5' to the start site of transcription. *J Immunol* 1993;151:2633–2645. [PubMed: 8360483]
114. Tee MK, Babalola GO, Aza-Blanc P, Speek M, Gitelman SE, Miller WL. A promoter within intron 35 of the human C4A gene initiates abundant adrenal-specific transcription of a 1 kb RNA: location of a cryptic CYP21 promoter element? *Hum Mol Genet* 1995;4:2109–2116. [PubMed: 8589688]
115. Wang Y, Macke JP, Merbs SL, et al. A locus control region adjacent to the human red and green visual pigment genes. *Neuron* 1992;9:429–440. [PubMed: 1524826]
116. Wurster AL, Siu G, Leiden JM, Hedrick SM. Elf-1 binds to a critical element in a second CD4 enhancer. *Mol Cell Biol* 1994;14:6452–6463. [PubMed: 7935370]
117. Fujiwara Y, Miwa M, Takahashi R, Hirabayashi M, Suzuki T, Ueda M. Position-independent and high-level expression of human alpha-lactalbumin in the milk of transgenic rats carrying a 210-kb YAC DNA. *Mol Reprod Dev* 1997;47:157–163. [PubMed: 9136116]
118. Raguz S, Hobbs C, Yague E, Ioannou PA, Walsh FS, Antoniou M. Muscle-specific locus control region activity associated with the human desmin gene. *Dev Biol* 1998;201:26–42.
119. Kamat A, Graves KH, Smith ME, Richardson JA, Mendelson CR. A500-bp region, approximately 40 kb upstream of the human CYP19 (aro-matase) gene, mediates placenta-specific expression in transgenic mice. *Proc Natl Acad Sci U S A* 1999;96:4575–4580. [PubMed: 10200304]
120. Heydemann A, Warming S, Clendenen C, Sigrist K, Hjorth JP, Simon MC. A minimal c-fes cassette directs myeloid-specific expression in transgenic mice. *Blood* 2000;96:3040–3048. [PubMed: 11049982]

121. Moon AM, Ley TJ. Conservation of the primary structure, organization, and function of the human and mouse beta-globin locus-activating regions. *Proc Natl Acad Sci U S A* 1990;87:7693–7697. [PubMed: 2217202]
122. Carson S, Wiles MV. Far upstream regions of class II MHC Ea are necessary for position-independent, copy-dependent expression of Ea transgene. *Nucleic Acids Res* 1993;21:2065–2072. [PubMed: 8502547]
123. Ganss R, Montoliu L, Monaghan AP, Schutz G. A cell-specific enhancer far upstream of the mouse tyrosinase gene confers high level and copy number-related expression in transgenic mice. *Embo J* 1994;13:3083–3093. [PubMed: 8039502]
124. Hammer RE, Krumlauf R, Camper SA, Brinster RL, Tilghman SM. Diversity of alpha-fetoprotein gene expression in mice is generated by a combination of separate enhancer elements. *Science* 1987;235:53–58. [PubMed: 2432657]
125. Jenuwein T, Forrester WC, Qiu RG, Grosschedl R. The immunoglobulin mu enhancer core establishes local factor access in nuclear chromatin independent of transcriptional stimulation. *Genes Dev* 1993;7:2016–2032. [PubMed: 8406005]
126. May G, Enver T. Targeting gene expression to haemopoietic stem cells: a chromatin-dependent upstream element mediates cell type-specific expression of the stem cell antigen CD34. *Embo J* 1995;14:564–574. [PubMed: 7532132]
127. Palmiter RD, Sandgren EP, Koeller DM, Brinster RL. Distal regulatory elements from the mouse metallothionein locus stimulate gene expression in transgenic mice. *Mol Cell Biol* 1993;13:5266–5275. [PubMed: 8355681]
128. Smith MS, Lechago J, Wines DR, MacDonald RJ, Hammer RE. Tissue-specific expression of kallikrein family transgenes in mice and rats. *DNA Cell Biol* 1992;11:345–358. [PubMed: 1605858]
129. Terajima M, Nemoto Y, Obinata M. Inducible expression of erythroid-specific mouse glycophorin gene is regulated by proximal elements and locus control region-like sequence. *J Biochem (Tokyo)* 1995;118:593–600. [PubMed: 8690723]
130. Elenich LA, Ford CS, Dunnick WA. The gamma 1 heavy chain gene includes all of the cis-acting elements necessary for expression of property regulated germ-line transcripts. *J Immunol* 1996;157:176–182. [PubMed: 8683112]
131. Sabbatini P, Georgiou A, Sinclair C, Dillon N. Analysis of mice with single and multiple copies of transgenes reveals a novel arrangement for the lambda5-VpreB1 locus control region. *Mol Cell Biol* 1999;19:671–679. [PubMed: 9858590]
132. Johnson BA, John VA, Henschler R, et al. Upstream elements bestow T-cell and haemopoietic progenitor-specific activity on the granzyme B promoter. *Gene* 1999;234:101–107. [PubMed: 10393244]
133. Arai Y, Kajihara S, Masuda J, et al. Position-independent, high-level, and correct regional expression of the rat aldolase C gene in the central nervous system of transgenic mice. *Eur J Biochem* 1994;221:253–260. [PubMed: 8168514]
134. Dale TC, Krnacik MJ, Schmidhauser C, Yang CL, Bissell MJ, Rosen JM. High-level expression of the rat whey acidic protein gene is mediated by elements in the promoter and 3' untranslated region. *Mol Cell Biol* 1992;12:905–914. [PubMed: 1545822]
135. Talbot D, Descombes P, Schibler U. The 5' flanking region of the rat LAP (C/EBP beta) gene can direct high-level, position-independent, copy number-dependent expression in multiple tissues in transgenic mice. *Nucleic Acids Res* 1994;22:756–766. [PubMed: 8139915]
136. Li QL, Zhou B, Powers P, Enver T, Stamatoyannopoulos G. Beta-globin locus activation regions: conservation of organization, structure, and function. *Proc Natl Acad Sci U S A* 1990;87:8207–8211. [PubMed: 2236034]
137. Slightom JL, Bock JH, Tagle DA, et al. The complete sequences of the galago and rabbit beta-globin locus control regions: extended sequence and functional conservation outside the cores of DNase hypersensitive sites. *Genomics* 1997;39:90–94. [PubMed: 9027490]
138. Whitelaw CB, Harris S, McClenaghan M, Simons JP, Clark AJ. Position-independent expression of the ovine beta-lactoglobulin gene in transgenic mice. *Biochem J* 1992;286(Pt 1):31–39. [PubMed: 1520282]

Figure 1. The human and mouse β -globin loci

The human locus consists of 5 functional genes, indicated as dark boxes, arrayed in their order of developmental expression, 5'- ϵ - $G\gamma$ - $A\gamma$ - δ - β -3'. There are 2 developmental switches in globin chain synthesis coincident with changes in site and type of erythropoiesis. During primitive erythropoiesis, the ϵ -globin gene is expressed in the embryonic yolk sac. The first switch occurs at approximately 8 weeks' gestation; the ϵ -globin gene is silenced and the $G\gamma$ - and $A\gamma$ -globin genes are expressed during definitive erythropoiesis in the fetal liver. The second switch occurs shortly after birth; the γ -globin genes are silenced and the β -globin gene and, to a lesser extent, the δ -globin gene are activated in the bone marrow. The HSs 5'HS1 through 5'HS7 are located -6, -11, -15, -18, -22, -28, and -35 kb relative to the ϵ -globin gene, respectively, and are indicated by arrows. 5'HS1 through 5'HS4 are erythroid specific, but 5'HS5 through 5'HS7 are not. Another HS (3'HS1) is located 20 kb downstream of the β -globin gene; 3'HS1 is found only in erythroid cells. Boxes represent globin genes and ovals represent olfactory receptor genes; filled ones represent the productive genes and shaded ones the pseudogenes. The lines below the diagram of the locus indicate deletions of the LCR discussed in "In vivo function of LCRs." The Hispanic deletion, which causes $(\gamma\delta\beta)^0$ thalassemia in humans, extends an additional 20 kb 5' of the LCR 5'HS5.

Figure 2. Models of LCR function

A globin gene is denoted by a green rectangular box with the promoter region indicated in a lighter green. Transcription factors are shown as colored ovals and circles. The 4 erythroid-specific hypersensitive site cores (HSs) are indicated by small red boxes. Blue boxes are the positions of 5'HS5 and 3'HS1, representing likely insulator elements. The flanking DNA sequences of the HSs are depicted as loops between the HS cores. Transcripts are denoted by wavy arrows. (A) Looping model. Transcription factors bind to the LCR HSs and the gene promoter. The LCR directly interacts with the gene promoter by looping out the intervening DNA, thus forming an active transcription complex at the gene promoter. (B) Tracking model. Sequence-specific transcription factors bind to the LCR, forming a complex that tracks down

the DNA sequence, as depicted by the large black arrowhead, until encountering transcription factors bound to the appropriate gene promoter, initiating high-level gene expression. (C) Facilitated tracking model. Aspects of both looping and tracking models are combined. Sequence-specific transcription factors bind the LCR; looping then occurs to deliver the bound transcription factors proximal to the gene promoter, followed by tracking, until they encounter transcription factors bound to the appropriate gene promoter. (D) Linking model. Sequential binding of transcription factors along the DNA directs changes in chromatin conformation and defines the transcriptional domain. The transcription factors are linked to one another from the LCR to the gene promoter by non-DNA-binding proteins and chromatin modifiers (shown as small colored circles).

Table 1**Mammalian LCR and LCR-like elements*****Human:***

- 1 β -globin locus¹
- 2 Adenosine deaminase gene, human⁸⁶
- 3 Apolipoprotein E/C-1 gene locus, human¹⁰⁵
- 4 T cell receptor α/δ locus, human⁹⁶
- 5 CD2 gene, human⁹⁵
- 6 S100 β gene, human¹⁰⁶
- 7 Growth hormone gene, human¹⁰⁷
- 8 Apolipoprotein B gene, human¹⁰⁸
- 9 β myosin heavy chain gene, human¹⁰⁹
- 10 MHC class I HLA-B7 gene, human⁸⁷
- 11 Immunoglobulin heavy chain locus, human¹¹⁰
- 12 Immunoglobulin C alpha 1 & 2 genes, human¹¹¹
- 13 Keratin 18 gene, human¹¹²
- 14 MHC class I HLA G gene, human¹¹³
- 15 Complement component C4A & B genes, human¹¹⁴
- 16 Red and green visual pigment genes, human¹¹⁵
- 17 CD4 gene, human¹¹⁶
- 18 α -lactalbumin, human¹¹⁷
- 19 Desmin gene, human¹¹⁸
- 20 CYP19 (aromatase) gene, human¹¹⁹
- 21 *c-fes* proto-oncogene, human¹²⁰

Mouse:

- 1 β -globin locus¹²¹
- 2 MHC Class II Ea gene, mouse¹²²
- 3 Tyrosinase gene, mouse¹²³
- 4 α -fetoprotein gene, mouse¹²⁴
- 5 Immunoglobulin μ heavy chain, mouse¹²⁵
- 6 CD34 gene, mouse¹²⁶
- 7 Metallothionein II gene, mouse¹²⁷
- 8 Kallikrein genes, mouse and rat¹²⁸
- 9 Glycophorin gene, mouse¹²⁹
- 10 γ 1 heavy chain gene, mouse¹³⁰

- 11 $\lambda 5-V_{\text{preB1}}$ locus, mouse¹³¹
- 12 granzyme B gene, mouse¹³²
- 13 T cell receptor γ locus, murine¹⁰²
- 14 Interleukin-2 gene, mouse⁸⁴

Rat:

- 1 Aldolase C gene, rat¹³³
- 2 Whey acidic protein (WAP) gene, rat¹³⁴
- 3 Kallikrein genes, mouse and rat¹²⁸
- 4 LAP (C/EBP β) gene, rat¹³⁵

Other:

- 1 β -globin locus, goat, rabbit^{136,137}
 - 2 β -lactoglobulin gene, ovine¹³⁸
-