THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Uncovering Enhancer Functions Using the -Globin Locus

OPEN ACCESS

PLOS | GENETICS

Review

Uncovering Enhancer Functions Using the α-Globin Locus

Douglas Vernimmen*

The Roslin Institute, Developmental Biology Division, University of Edinburgh, Easter Bush, Midlothian, United Kingdom

Abstract: Over the last three decades, studies of the α- and β-globin genes clusters have led to elucidation of the general principles of mammalian gene regulation, such as RNA stability, termination of transcription, and, more importantly, the identification of remote regulatory elements. More recently, detailed studies of α-globin regulation, using both mouse and human loci, allowed the dissection of the sequential order in which transcription factors are recruited to the locus during lineage specification. These studies demonstrated the importance of the remote regulatory elements in the recruitment of RNA polymerase II (PolII) together with their role in the generation of intrachromosomal loops within the locus and the removal of polycomb complexes during differentiation. The multiple roles attributed to remote regulatory elements that have emerged from these studies will be discussed.

Introduction

Developmentally regulated genes become activated only in the appropriate lineage, while they remain inactivated (or become fully repressed) in other lineages. Haematopoiesis offers an excellent system to study gene regulation because it is very accessible and produces a palette of a least eight different lineages in the blood [1]. Also, differentiated cells and their progenitors can be easily separated and purified using different cell surface markers [2,3]. The erythroid cell lineage is of particular interest because of the high expression level of the globin genes, which produce the major proteins found in red blood cells. Adult haemoglobin (HbA) is made by the formation of a tetramer containing two α-globin chains and two β-globin chains. The level of expression of these proteins needs to be equimolar to ensure the correct formation of this tetramer; an imbalance would create insoluble homotetramers, the key pathological feature of thalassaemia [4].

The expression of adult α- and β-globin genes requires a panel of different tissue-specific transcription factors (TFs), including GATA1, GATA2, nuclear factor-erythroid 2 (NF-E2), stem cell leukaemia factor (SCL), and Krüppel-like factor 1 (KLF1) (formerly called Erythroid Krüppel-like factor [EKLF]). These TFs are expressed at different times during differentiation, suggesting specific roles for each [5]. For example, GATA2 is expressed early, in common myeloid progenitors, whereas GATA1 is expressed later, in erythroid progenitors. KLF1 is also expressed late during erythroid differentiation and therefore is important for late events in erythropoiesis such as the expression of the α- and β-globin genes [6,7].

Enhancers were originally defined as sequences that increase the rate of transcription of a target gene [8–10]. They may lie far away upstream or downstream from the gene they regulate [10] and should work in both orientations [11]. These original definitions were based on reporter assays (i.e., plasmids) in which the distance separating enhancers and promoters is very small and the chromatin context is not taken into account. The activity in both orientations was primarily due to the fact that simple enhancers were identified as regions formed by a palindromic sequence (e.g., binding site of a homodimer [12]). It became more obvious later that enhancers usually work in groups (i.e., locus control region (LCR) and super enhancers [13]), each being bound by several TFs, forming a so-called enhanceosome [14,15]. These enhanceosomes are nucleated by pioneer TFs early during differentiation and subsequently are replaced by other TFs that trigger transcription by PolII recruitment. More recently, in vivo studies found that enhancers (remote regulatory elements) can be located sometimes up to 1 Mb away from the gene they regulate [16–18]. What do these sequences do? How do they function across very large distances? These have been questions of major interest over the last two decades. Studies on the α- or β-globin loci have been pioneers in this field, and many approaches and tools have been developed to address this.

Overall, three main strategies have been used to understand the molecular mechanisms involved in the activation of the globin genes. The first involves biochemistry and molecular biology using different cell types representing different stages of differentiation and the analysis of TF complexes, TF binding patterns, chromosome conformation, and epigenetic changes. The second involves mouse genetics, and the third strategy involves the study of patients carrying different types of mutations (in these cases deletions) associated with the downregulation of expression of globin genes leading to thalassaemia. This review will be focused on studies made on both the mouse and the human α-globin loci and more particularly on the role attributed to the remote regulatory elements controlling their expression.

Structure of the α-Globin Locus

Studies of both the mouse and the human α-globin loci offer complementary advantages. The mouse locus can be manipulated easily, and the repertoire of cell lines is greater, whereas the human locus can be studied in samples obtained from patients [4]. The α-globin locus of all mammalian species analysed lies within a region of 135–155 kb of conserved synteny, with the α-like genes arranged along the chromosome in the order 5′-ζ-α-α-3′ (Figure 1) [19]. However, the mouse and human loci also show some important structural differences. The mouse locus has an arrangement containing two pseudogenes θ (5′-ζ-α1-5′θ-α2-3′θ-3′, Figure 1), whereas the human has an arrangement containing only one pseudogene θ, with another pseudogene (αD) that precedes the adult α genes (5′-ζ-αD-α2-α1-θ-3′). The αD gene is expressed in erythroid cells [20]. Also, the human α-like genes are each covered by a CpG island (Figure 1), which has a strong influence on their regulation (see below). The erythroid-specific multispecies conserved sequences (MCS) identified by DNase hypersensitive sites (DHS) have been numbered MCS-R1 to MCS-R4 [19] (Figure 1). Three of these elements (MCS-R1, MCS-R2, and MCS-R3) lie within the body of a housekeeping gene, NPRL3, and MCS-R4 lies upstream of the promoter of that gene (Figure 1). However, the mouse locus contains an additional DHS, located 12 kb upstream of the ζ gene (HS –12), that appears

to be important during the priming of the locus during differentiation [5].

The globin gene disorders (haemoglobinopathies), including thalassaemias, are among the most common human genetic diseases, with more than 300,000 severely affected individuals born throughout the world every year [21]. Thalassaemias are characterized by inherited mutations leading to a reduction of the synthesis of α- (α-thalassaemia) or β-globin (β-thalassaemia) chains from one allele. Human genetics is a good approach to identify functional remote regulatory elements, and original observations were made in patients with α- and β-thalassaemia [22–24]. In most cases, a deletion removing a globin gene is the cause of this down-regulation, but in some rare cases, the genes (including their promoters) remain intact [25].

In rare cases of α-thalassaemia, further chromosome mapping in a number of patients led to the identification of other deletions located far upstream, overlapping the NPRL3 housekeeping gene [26] and therefore the remote erythroid hypersensitive sites (Figure 1). By comparing all the different deletions in the NPRL3 gene, it appears that MCS-R1 and MCS-R2 are consistently removed in all cases, and this led to the characterisation of these sequences [27,28]. From these observations, different approaches, such as transient transfections and transgenics, have showed that MCS-R2 (previously called HS –40) has a much stronger enhancer activity. Moreover, this element shows a remarkable
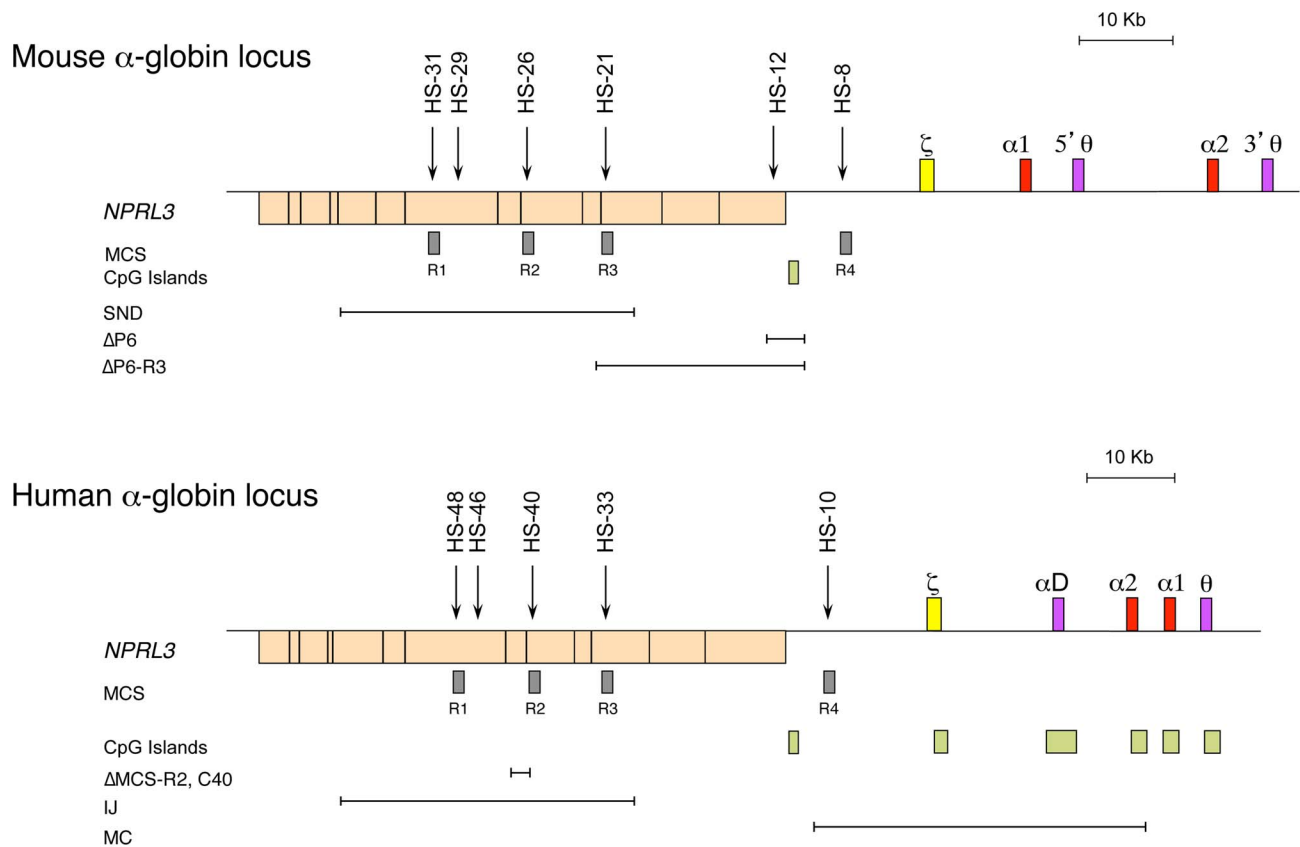
Figure 1. The chromosomal organisation of the mouse (top) and human (bottom) α-globin clusters. The embryonic ζ gene in each locus is represented in yellow, the pseudogenes in purple, and the foetal/adult α genes in red boxes. The positions of DNase I hypersensitive sites, discussed in the text, are shown as arrows. The widely expressed gene NPRL3 is transcribed from the opposite strand to that of α-globin and is shown as an orange box, with exons in black bars. Grey boxes refer to previously defined multispecies conserved elements (MCS), and light-green boxes indicate CpG islands. Note the lack of CpG islands in the mouse locus. The previously described natural (IJ, MC) and targeted (C40, ΔMCS-R2) deletions from the human and targeted deletions from the mouse (ΔP6, ΔP6-R3, and SND) α-globin cluster are shown as annotated.
doi:10.1371/journal.pgen.1004668.g001

conservation throughout evolution. In zebrafish, it already lies in the intron of the *NPRL3* gene and drives the expression of both α- and β-globin on the same locus through a bidirectional promoter, securing comparable levels of α- and β-like globin protein [29]. The globin locus in fish has then diverged over time and segregated into separate α and β loci after the divergence of amphibians 450 million years ago, leaving the ancestral globin locus as an α-globin locus [30,31]. Therefore, further efforts were concentrated on the characterisation of the ancestral MCS-R2 element in detail [28,32–34].

Epigenetic Control of α-globin Gene Expression

The epigenetic programme seems to play a key role in determining cell fate, including the decision to undergo self-renewal or commitment. Chromatin immunoprecipitation (ChIP) followed by high-throughput sequencing (ChIP-Seq) studies have suggested that the chromatin associated with many genes controlling lineage fate decisions is uniquely marked in stem cells [35]. Their histone signature includes modifications associated with both transcriptional repression (H3K27me3) imposed by the polycomb group proteins (PcG) and activation (H3K4me3) mediated by the trithorax group proteins (TrxG) [36]. The "permissive" chromatin state of these genes in stem cells is called bivalent and is supposed to be resolved during differentiation, as genes become fully activated or repressed [37–39]. The recruitment of complexes such as PcG and TrxG involves the presence of a CpG island at the target promoter of a developmentally regulated gene, and proteins with a CXXC domain binding to the DNA [40–44]. The maintenance and propagation of an epigenetic mark, such as H3K27me3 by polycomb repressive complex 2 (PRC2), is well documented and involves a "reader" protein (EDD), which recognises a modified histone, and a "writer" protein (histone methyltransferase Ezh2), which modifies the histones nearby [45]. The removal of such a mark by an "eraser" protein (histone demethylases JMJD3 and UTX) would prevent the maintenance and propagation from occurring. The human α-globin genes are associated with a CpG island, whereas in rodents (e.g., mouse and rat) this CpG island has been lost during evolution (Figure 1) [43]. Therefore, the presence of a CpG island on the human locus has an important implication in the epigenetic regulation of that locus (see below).

In mouse embryonic stem (ES) cells, the analysis of the mouse α-globin locus initially suggested it to be unmarked by any histone modifications such as H4ac or H3K4me2 [5]. More recent genome-wide studies (ChIP Seq) have shown that in ES cells, remote regulatory elements of a small number of developmentally regulated genes are already marked by H3K4me1, and in the mouse α-globin locus, a deposition of H3K4me1 has been found just next to MCS-R1 (HS –31) in ES cells [46], suggesting a possible priming around that region.

In human ES cells and in nonerythroid cells, the CpG island is associated with PRC2 recruitment and its associated signature (H3K27me3), whereas these are not found in the rodent α-globin genes [47]. In human ES cells, the α-globin CpG islands are also covered by H3K4me3. However, the domains of H3K27me3 and H3K4me3 do not entirely overlap, and the ratio observed progressively changes during differentiation [48], suggesting that the progressive increase of H3K4me3 could be due to an expression of the gene at irrelevant stages rather than a signature of priming, as suggested by previous studies [36,49]. Therefore, single-cell mRNA expression in human ES cells was measured and revealed that although full length α-globin mRNA was detected in a small number of Oct4 positive cells, this number increased

during differentiation. This suggested that some bivalent domains might be the consequence of a subpopulation phenomenon rather than true priming in all cells [48]. PcG complexes remain until the last stages of differentiation [48].

Locus Priming during Differentiation

Developmentally regulated genes are progressively primed during the differentiation programme, a process that involves the sequential recruitment of stage-specific TFs associated with histone modifications. The study of α-globin was initiated using the mouse locus and a panel of different cell lines and primary cells that represent the different stages of erythroid differentiation (Figure 2) [5]. As a source of multipotent cells, mouse ES cells have been used, and so far no TFs have been found bound within the mouse α-globin MCS-Rs. This contrasts with other single-locus studies such as the mouse β-globin locus [50] and the λ5-VpreB1 locus [51] showing priming at enhancers (including PolII recruitment) in ES cells, long before transcription occurs (differentiated erythroid and B cells, respectively).

As a source of common myeloid progenitor cells (CMP), factor dependent cell Patterson (FDCP)–mix cells have been used [52]. These are able to produce both megakaryocytic and erythroid lineages. In these cells, the mouse α-globin locus is primed at MCS-R2 and HS –12 by GATA2 (Figure 2B) [5]. During differentiation, GATA2 regulates the expression of GATA1 with a negative feedback loop, which shuts down GATA2 expression in proerythroblasts [53]. As a source of proerythroblasts, the MEL (mouse erythroleukemia) cell line has been used. This cell line was originally transformed by the Friend virus [54] and is a very useful model to study the last stages of erythroid differentiation since, like most blood cell lines, it can be induced towards terminal differentiation after exposure to certain chemicals. Under treatment with inducing agents, such as dimethyl sulphoxide (DMSO) or hexa-methylene bis-acetamide (HMBA), MEL cells turn a red colour because of haemoglobinisation. At the proerythroblast stage (uninduced MEL), most remote regulatory elements are formed (HS –12, MCS-R1, MCS-R2, and MCS-R4) with the recruitment of key erythroid TFs, including GATA1, NF-E2 (dimer p45 and p18), and the SCL pentameric complex (Figure 2C) [5]. The promoter also becomes accessible and bound by GATA1, 89-kDa zinc finger protein (ZBP-89), and nuclear factor Y (NFY) (CAAT-box-binding protein). The human α-globin promoter does not contain an evolutionarily conserved GATA site and does not bind ZBP-89, but NFY binding is common in both species [7]. NFY has a histone-like structure [55] and therefore could be involved in chromatin opening by nucleosome replacement [56]. In erythroblasts (induced MEL), more changes occur, with the appearance of MCS-R3 (HS –21), together with recruitment of KLF1 (Figure 2D) and the preinitiation complex (PIC) at both enhancers and promoters in the human and the mouse α-globin loci (Figure 2E) [7]. At that final stage, other Sp/X-Kruppel-like transcription factors (Sp/X-KLFs) are also recruited to the α-globin promoters [57,7]. Interestingly, a looped structure that bridges the remote regulatory elements to the promoter was also detected by chromosome conformation capture analysis using quantitative Taqman technology (q3C) at that particular stage of differentiation (Figure 2F) [7], where high level of α-globin expression occurs. *NPRL3* gene expression is also up-regulated in human [58] and mouse [59] erythroid cells and associated with bidirectional intergenic transcripts at these elements (enhancer RNA [eRNA]) [59]. The more recent advances of ChIP-Seq technologies allowed other groups to find the same phenomenon with peaks of PolII in intergenic regions [60–62] associated with
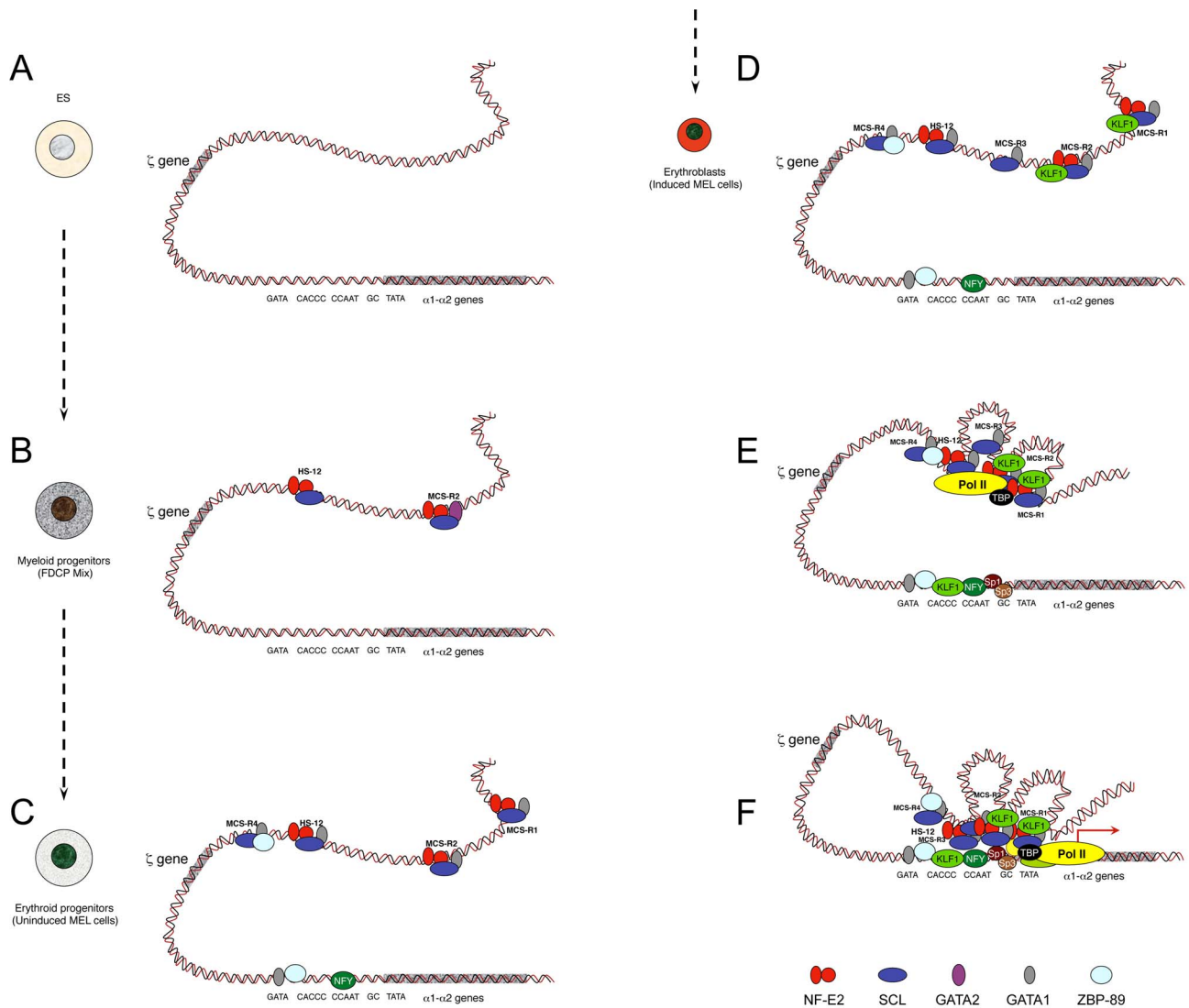
Figure 2. Model showing sequential order of TF binding on the mouse α-globin locus during erythropoiesis. The locus has been analysed in pluripotent cells (A), myeloid progenitors (B), committed erythroid progenitors (C), and in differentiating erythroid cells (D, E, F). Note that D, E, and F do not represent different steps during activation but simply separate illustrations representing the formation of MCS-R3 and recruitment of KLF1 (D), the recruitment of PolII at the enhancers independently of the promoter (E), and a multilooped structure involving enhancers and promoters (F). Note that the human α-globin locus has been mainly characterised in human ES cells and primary erythroid cells. It is still not known if MCS-R2 is also primed in human progenitor cells (i.e., myeloid progenitors) as it is for the mouse locus. Studies have shown that the TF binding pattern is essentially the same with two main key differences: GATA1 does not bind the human promoter in erythroid cells, and the gene is repressed by PcG complexes in human ES cells. PcG complexes are removed late during erythropoiesis. For the sake of clarity, the locus has been represented as a double helix of DNA rather than by a 10 nm chromatin fibre.
doi:10.1371/journal.pgen.1004668.g002

H3K4me1, a key feature of enhancers [63]. Note that the characterization of the human α-globin locus in progenitor cells has not been characterized to the same depth as in differentiated cells because the repertoire of progenitor cells is less accessible. Indeed, progenitor cells can be obtained from bone marrow samples from individuals undergoing total hip replacement for osteoarthritis [48] or after mobilisation for bone marrow transplant, but the amount of cells collected is still very limited.

Role(s) of the Remote Regulatory Sequences

The natural deletions observed in patients were an essential tool in the analysis of the role of these remote regulatory elements. Due to the nature of the disease (anaemia), it was however difficult to

perform a number of experiments on primary cells because of the large amount of material required and also because this material is only available for a limited time. For this reason, interspecific hybrids were produced by the fusion of a human immortalised B cell (Epstein-Barr virus [EBV] infected) with the mouse MEL cell line described earlier [24,28,64,65].

Therefore, interspecific hybrids, derived from normal individuals or those with previously characterised natural mutations of the α-globin cluster, were analysed to determine if the PIC was recruited independently both at the promoter and at the remote regulatory elements. Two types of mutation were analysed: one in which the remote upstream elements had been fully (IJ) deleted but the α promoters remained intact [27,66] and another in which all α-like genes were deleted (MC) but the upstream elements were

still present (Figure 1). From these studies, two important principles in the hierarchy governing the process of transcription activation have been uncovered. First, the presence of the remote enhancers is required for the recruitment of the PolII at the promoter. Secondly, the presence of the promoter is not required for PolII recruitment at the enhancers [7]. In other words, the recruitment of PolII at the enhancers occurs independently of the promoter. In a similar way, preventing PolII recruitment at the mouse β-globin promoter does not affect its binding at the upstream LCR [67]. This important feature was validated later on a study on the *Arc* locus [61] where, in the *Arc* gene knockout neurons, PolII remains bound at the Arc enhancer at levels equivalent to those observed in wild-type neurons. In the former study, eRNA synthesis is abolished in the absence of the gene, suggesting that, like mRNA synthesis, eRNA synthesis may require an interaction of the enhancer with a promoter [61]. However, another study on the human growth hormone (*hGH-N*) locus showed that this is not always the case [68].

The use of interspecific hybrids is a useful way to study enhancer functions; however, they are very difficult to maintain, mainly because they tend to lose the human chromosome 16 and they are still also very difficult to induce into erythroblasts. More recently, another mouse model was generated in which the whole region of conserved synteny (87 kb of mouse genomic DNA) was replaced by the human orthologous region (117 kb of human genomic DNA) [69]. Therefore, a humanized mouse was produced, which offers the advantage to avoid the issues generated by the use of conventional transgenics [69,70]. Although the spatiotemporal expression of the human α-globin was correct, the level seemed to be suboptimal at about 50% of what was expected. However, in this model, the pattern of TF binding (including PolII) and the looped structure seem to be identical to that in human primary erythroblasts [71,72]. This new model was therefore the best one to use so far to investigate the role of MCS-R2. A deletion of about 1.1 kb covering MCS-R2 was made in heterozygous mice (one humanised chromosome and one normal mouse chromosome). In this situation, the level of α-globin expression reached only ∼2% compared to the normal humanised chromosome [69,71]. These data recapitulated well the previous data obtained with other models, including a hybrid clone bearing the same deletion (Figure 1, C40 hybrid [27]).

The changes of TF binding and chromosomal looping were then analysed in normal humanised mice and compared to humanized mice without MCS-R2 (ΔMCS-R2) [73]. In the ΔMCS-R2 mutant, the absence of α-globin expression is associated with a lack of PolII binding together with an impaired looping formation. The occupancy of PolII from the start to the end of the gene is about 6%–3% of normal, matching the amount of mRNA expressed, and therefore ruled out any role of MCS-R2 in PolII elongation. A role in PolII elongation was originally suggested for the LCR of the β-globin locus [74]. In this study, a targeted deletion of the human LCR in transgenic mice reduces PIC recruitment of about 50% to normal but with a more dramatic effect on Ser5 phosphorylation of PolII (Ser-5P) and transcriptional elongation, suggesting a role for the β-globin LCR in both PolII recruitment (partially) and elongation. Similar observations were made later with the mouse β-globin locus, in which a targeted deletion of the mouse LCR reduces global PolII recruitment at the promoter to 30% of normal with a redistribution of Ser-2P along the gene [75]. Recently, the Furlong group showed that in *Drosophila* enhancers would loop towards a target promoter to recruit PolII but paused in the majority of cases. Releasing of PolII pausing would then occur by subsequent recruitment of additional TFs or additional enhancers [76]. Note

that in the studies of the α-globin locus, kinases regulating PolII such as cdk7 (TFIIH), cdk-8 (mediator), and cdk9 (elongation factor p-TEFb) are all recruited at the same time on the promoter when looping and transcription occur [7]. On the mouse β-globin locus, cdk9 and PolII Ser-2P binding is observed at both the LCR and the promoter when the gene is active [75]. Although the binding of cdk9 to the LCR precedes its binding to the promoter, the deletion of the LCR does not affect cdk9 binding at the promoter. This study would suggest that the β-globin LCR may not control elongation by delivery of cdk9 but would involve other complexes such as DRB sensitivity inducing factor (DSIF) and facilitates chromatin transcription (FACT) [77]. Another study on the Myb locus in erythroid cells [78] together with a genome-wide study in human B lymphoblast cells (MM1.S) [79] suggested an important role of cdk9 bound to enhancers. Other genome-wide studies detected one (Ser-5P) [60] or both (Ser-2P and Ser-5P) [80] phosphorylated forms of PolII at active enhancers, supporting the idea that enhancers would deliver an activated PolII at the target promoters. A general role of remote enhancers in PolII delivery to a target promoter has been also described in other loci such as the human growth hormone (*hGH*) locus [81,82], the human serpin cluster [83], and the mouse T cell receptor beta (TCRβ) locus [84].

In the ΔMCS-R2 mutant, the PolII occupancy at the other remaining enhancers is also affected but not as much as at the promoter. In fact, the MCS-R1 element, located upstream of the deletion, is less affected (∼20% of normal PolII occupancy) than the other regions downstream (MCS-R3, MCS-R4, and the promoter). The same trend is observed with other components of the PIC and other TFs such as KLF1 (Figure 3) [73]. Strikingly, this same observation was previously made with the C40 hybrid, albeit with a less dramatic reduction of TF binding (general transcription factors [GTFs] are unchanged at MCS-R1, and PolII occupancy level at the promoter is of about 20% compared to normal; [7], see discussion below). This reduction of TF occupancy across the locus suggests that the activating signal is propagating from the upstream elements towards the downstream promoter. The nature of this signal polarity is not yet understood and still has to be tested experimentally. Note, a polarity has been shown with the β-globin LCR locus, which works in an orientation-dependent manner [85]. Although this could also suggest a form of facilitated tracking, involving a unidirectional (multi)looped structure progressing towards the downstream promoters [86–89], there was no evidence of TF binding in the intervening DNA regions. Keeping this idea in mind, MCS-R2 was reinserted in an ectopic site in the region of conserved synteny, just downstream from the α-globin genes (Figure 4) [71]. In this new mutant (3′ MCS-R2), transcription is reactivated to about ∼50%, and more importantly, the looped structure is reestablished. This would argue against a facilitated-tracking model since, in this situation, the other remote elements (MCS-R1, MCS-R3, and MCS-R4) interact again with MCS-R2 and the α-globin genes, both located on the other side of the loop (Figure 4).

The analysis of MCS-R2 was completed by investigating its potential role in the deposition and/or removal of histone modifications, which are associated with activation and repression. Since the literature has reported about 150 possible histone modifications [90], this α-globin study concentrated on the main modifications associated with activation (H3 and H4 acetylation, H2B ubiquitination, H3K79me3, H3K4me3, and H3K36me3) and repression (H3K27me3). Repression associated with polycomb repressive complex 1 (PRC1) was not investigated because of the poor efficiency of the antibodies against PRC1 proteins and their associated modification (H2Aub). In agreement with the very
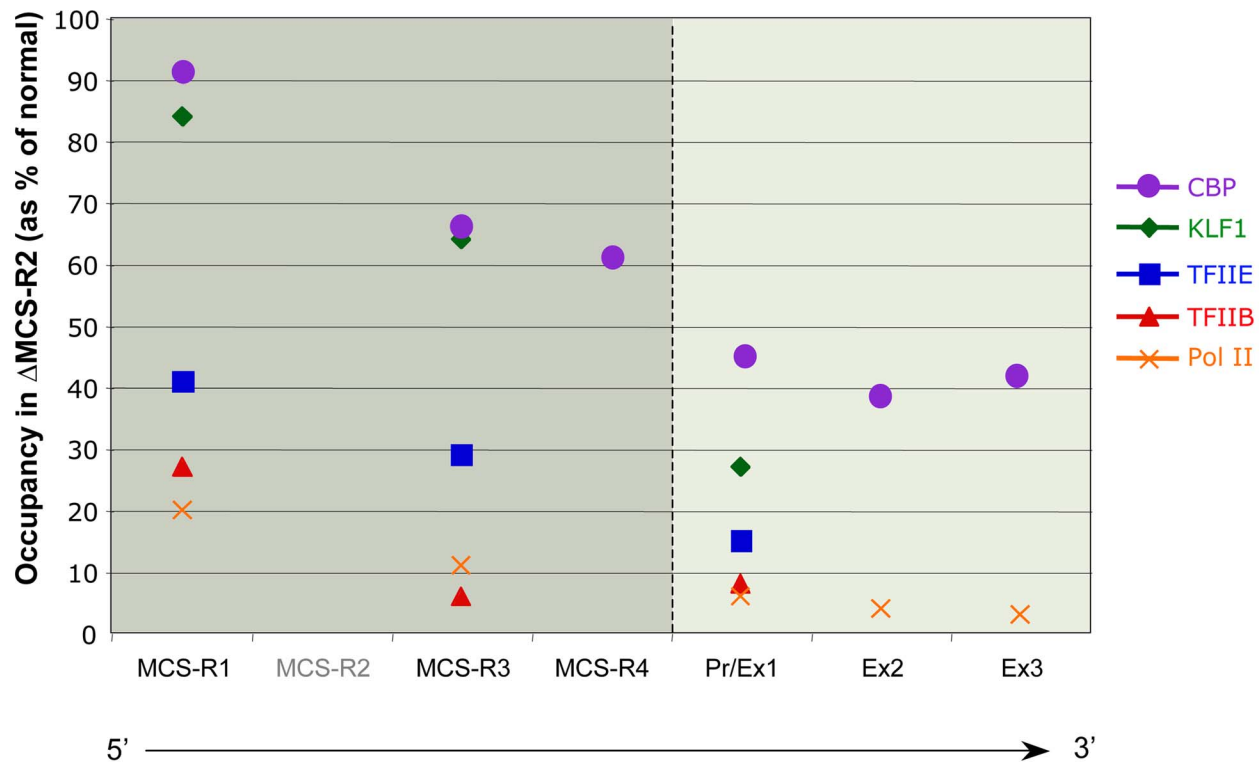
Figure 3. Differential occupancy of TFs in the absence of the enhancer (MCS-R2) in mouse primary erythroblasts. The percentages of occupancy in the mutant human allele (ΔMCS-R2) was calculated by comparison with the normal human allele (=100%) after normalising the ChIP efficiency to an endogenous control (mouse α-globin allele). This graph shows a reduction of TFs occupancy (5′→3′) with the distance along the locus, with the lowest occupancy observed at the gene (Pr/Ex1 – Ex2 – Ex3). Figure adapted from Vernimmen et al., 2011; Genes and Development [73].
doi:10.1371/journal.pgen.1004668.g003

low expression of the gene without MCS-R2, modifications such as H2Bub and H3K79me3 do not occur. Histone 3 mono- and di-methyl K79 are less affected by the deletion and are also found at the other remote regulatory sequences, whereas H3K79me3 was exclusively found at the body of the gene. Surprisingly, all modifications generated by the TrxG pathway (H3K4me1, H3K4me2 and H3K4me3) occur normally in the absence of MCS-R2. Histone 3 tri-methyl K4 is, however, suboptimal at the
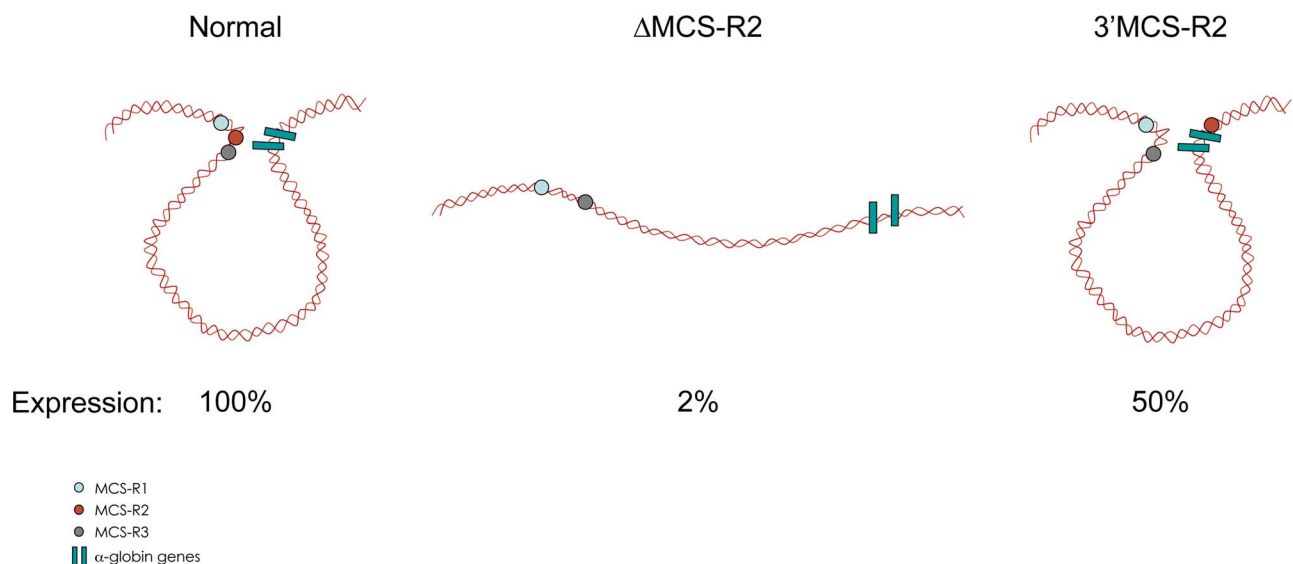

Figure 4. Communications between enhancers influences transcription. Intrachromosomal interactions involving human remote regulatory sequences and the human α-globin genes in normal, ΔMCS-R2, and 3′MCS-R2 humanised alleles in mouse primary erythroblasts.
doi:10.1371/journal.pgen.1004668.g004

body of the gene, and H3K4me1 is not only detected at the other MCS-R but is also significantly enriched at the body of the gene. It is worth noting that these are both common features observed in ES cells [48].

It was originally thought that the generation of H3K4me3 in ΔMCS-R2 mutant is due to the presence of the CpG island, since it was previously reported that the α-globin genes were already covered by this activating mark in human ES cells [47]. Moreover, the work from Adrian Bird's group has shown that a CXXC domain protein, cfp1 (also called CpG binding protein [CGBP]) is involved in the deposition of H3K4me3 at the CpG island of target genes regardless of their transcription status [44]. CGBP is part of a TrxG complex (human Set1) and therefore would generate H3K4me3 at any CpG island. In ChIP analysis, CGBP is only detected at the α-globin CpG island in the presence of MCS-R2 [73], suggesting another mechanism for H3K4me3 deposition in ΔMCS-R2 mutant. It was also reported that the binding of CGBP could be mutually exclusive with other proteins binding to CpG islands such as PcG [44]. Indeed, in the absence of MCS-R2, PRC2 (including histone deacetylase 1 [HDAC1]) recruitment continues at the α-globin CpG island throughout the whole process of erythroid differentiation, whereas this complex is removed in the presence of MCS-R2 enhancer (Figure 5). Importantly, the removal of PRC2 and H3K27me3 was associated with the recruitment of the H3K27me3 demethylase JMJD3 at the CpG island [73].

Conclusions

Using α-globin as a model, the study of enhancer biology was made possible by using cell lines and primary cells that faithfully represent the different stages of erythropoiesis. The sequential order of TF binding, with the appearance of DHSs, is followed by the recruitment of PolII at both enhancers and promoter at the last stages of erythropoiesis, when globin expression occurs. This contrasts with studies on the mouse β-globin and the λ5-VpreB1 loci in which PolII is also recruited late at the promoter but is detected very early at the remote regulatory elements [50,51]. Using human genetics, it has been possible to show that the recruitment of PolII at the remote enhancers occurs independently of the promoter. The general idea that enhancers act as docking sites for the recruitment of the general transcription machinery (components of the PIC and not just PolII) has been suggested in a few studies [60,91,92].

Once the PIC is recruited to the enhancer, this transcription machinery would be then somehow transferred to the downstream promoter by a looping mechanism. Note that although the duplicated α-globin genes in most species have similar or identical
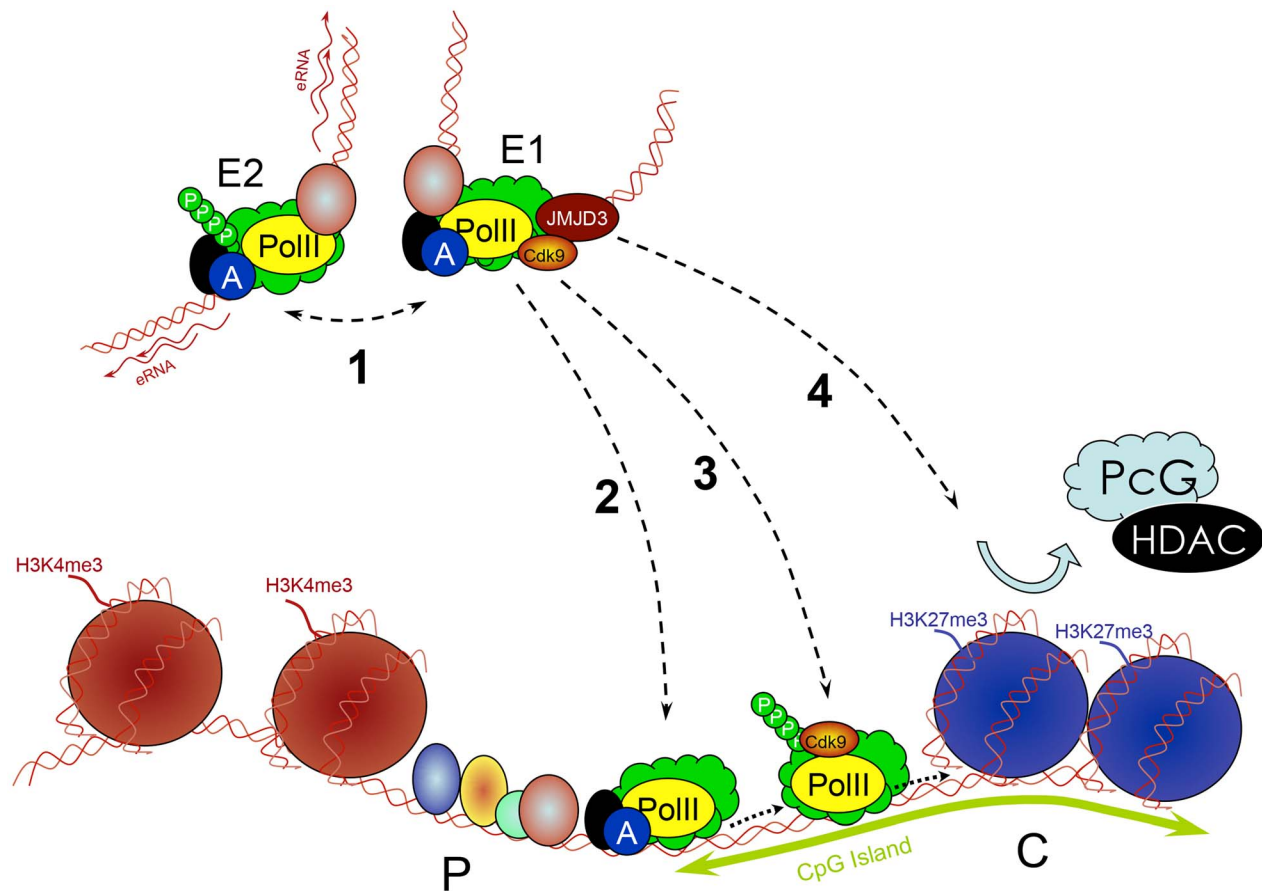
Figure 5. Multiple roles of enhancers. Enhancers (E) work in groups. They recruit PIC and TFs and interact with each other (1), which is associated with increased PolII concentration and production of eRNA. Enhancers subsequently deliver the PIC to a target promoter (P) through a looped structure (2). They control transcription elongation of some genes (e.g., by delivering an important kinase such as cdk9) by increasing the phosphorylation of PolII (3). They remove PcG repressive complexes from CpG islands of developmentally regulated genes through a histone demethylase (e.g., JMJD3) removing H3K27me3 (4).
doi:10.1371/journal.pgen.1004668.g005

promoters, the gene closest to the upstream elements is usually expressed at the higher level (reviewed in [93]). When more than two genes are present, the additional genes lying downstream appear to be expressed at even further decreased levels [94] and thus suggest the possibility of a local tracking process involved. However, this would be unlikely since the terminator element of the first gene should block this tracking [81], and the second gene (in the mouse) is flanked by CTCF [95], which should also block such a tracking mechanism [87]. By using both samples from patients and humanised mice bearing a deletion covering the major remote element (MCS-R2), it has been possible to show that the enhancer is required for the recruitment of the PIC and key TFs at the promoter. How the PIC is recruited to enhancers is still not known. An in vitro study on the β-globin locus suggested that TFs bound to enhancers are able to recruit directly the PIC [96], but in vivo studies on the same locus suggest that the β-globin LCR facilitates the localisation of the locus in active transcription sites termed transcription factories [77]. Both the α- and β-globin loci are associated with active transcription sites in the nucleus of living cells [97], but it is still left to debate if these transcription factories are simply reflecting focal accumulations of PolII on transcribing genes or preassembled organising structures to which genes move (reviewed in [98]).

Developmentally regulated genes are controlled by several enhancers, recently termed super enhancers [99], and q3C analysis showed that α-globin enhancers also communicate with each other (Figure 2E and Figure 4) [71]. These enhancers are primed at different times during differentiation, and this might create a sort of directionality of the signal. Interestingly, the deletion of MCS-R2 creates a reduction of TF binding downstream, across the locus (Figure 3) [73]. This could be the result of an impaired multilooped structure involving all the enhancers and the two α-globin promoters together. What ties the loop on the α-globin locus still has to be determined. Many proteins involved in the interactions between enhancers and promoters have been described (reviewed in [100]), and the role of each on the α-globin locus still has to be investigated.

CpG islands were originally found in housekeeping genes and are involved in keeping the chromatin open and therefore allowing basic transcription to occur [101]. However, we now know that CpG islands are also found in the promoters of developmentally regulated genes, which are not always active. For this group of CpG island promoters, PcG act to block the promoter accessibility in inappropriate lineages or at other stages during the differentiation programme. Nevertheless, it is not clear why developmentally regulated genes should have CpG islands in the first instance. In stem cells, the level of H3K4me3 observed at the human α-globin gene is indeed associated with full-length transcripts, albeit at a very low level [48]. The mouse α globin seems to be perfectly regulated in the absence of a CpG island. Although the level of full-length transcripts increases together with H3K4me3 during differentiation, H3K27me3 and PcG complexes seem to remain until the last stage of differentiation. The humanised mouse model made it possible to show that the removal of this repressive mark was dependent on the MCS-R2 enhancer, together with the recruitment of demetylase JMJD3 [73]. Of interest, another mutant with MCS-R2 deletion (C40 hybrid) did not show any persistence of PcG at the α-globin gene (unpublished data). This deletion, however, was made in a hybrid cell line [27], corresponding to the proerythroblast stage, which was described earlier (Figure 1). Thus, depending on the timing when the deletion was made (ES cells, i.e., prior to development versus proerythroblats, i.e., after differentiation), different results can be found [102]. Interestingly, the level of transcription in the absence

of MCS-R2 was similar, and therefore the increased level of transcription might not be the cause of PcG eviction. Also, the fact that PcG eviction is associated with the recruitment of a H3K27me3 demethylase suggests that it is more likely to be an active process [73]. It is therefore possible that enhancers could also recruit enzymes that remove histone modifications generated by PcG complexes at the target promoter after chromosomal looping, as it was suggested for the PIC. JMJD3 enrichment is not detected at MCS-R2 in humanized mouse erythroid cells by ChIP-qPCR, although this is at a much lower resolution [73]. However, ChIP-Seq studies showed enrichment of JMJD3 at both genes and distal intergenic regions [103,104]. The role of remote enhancers in the eviction of repressive PRC2 and PRC1 is also supported by other recent studies [105–107].

This all together led to the conclusion that MCS-R2 has multiple roles, and these may be applicable to any other enhancer: recruitment of PolII and key TFs at the promoter, formation of a looped structure involving several remote regulatory elements, and the removal of repressive complexes such as PcG (Figure 5). It thus seems that enhancers have evolved with the overall complexity of mammalian transcription regulation by developing multiple roles required for optimal gene expression. These different roles might be allocated to different enhancers controlling the same gene; some might be important for initiation of gene expression, others for its maintenance. How to disentangle these roles remains a challenge, since most of the time many events occur at the same time during differentiation. Alternative models allowing short kinetics (e.g., lipopolysaccharide (LPS)-activated macrophages [108]) might be more suitable for the study of gene activation when applicable. Indeed, the latest reflects gene activation in a given cell type that is under a physiologic change, whereas globin genes activation occurs with a change of cell identity during differentiation.

Future Directions

Over the last five years, high-throughput sequencing technology has been used to characterise human genetic diseases using materials from patients but also has provided a very broad picture of the chromatin landscape in many cell types. However, functional analyses are still required to provide mechanistic insights about enhancer functions. Over the last 30 years, single-locus dissection such as what has been achieved at the α- and β-globin loci remains a good example of where human genetics and genetic engineering in the mouse can provide important answers on how mammalian genes are regulated during differentiation and development. This is also applicable to other loci. For example, studies on patients with preaxial polydactyly led to the identification of the regulatory element controlling the expression of sonic hedgehog (SHH) in the posterior part of the limb [109]. Identifying individuals with α-thalassaemia by regular screenings allowed the mapping of all the known mutations, but sometimes new rare variants can be found, and these can provide new mechanisms underlying human genetic disease [21]. Human genetics, although powerful, still adds a layer of complexity when genetic variability between individuals has to be taken into account. Recently, a few rare cases of patients homozygous for MCS-R2 deletion have been described [110,111]. Surprisingly, these patients present a phenotype less severe than expected, even with a broader deletion removing the other remote regulatory elements on one chromosome [110]. This demonstrates that the other regulatory elements must have a role but also suggests that, when deleted, "orphan" enhancers might relay for α-globin expression as recently suggested for other genes (enhancer

adoption [112]). It is worth noting that in humans another enhancer-like element has been found outside the region of conserved synteny, 400 kb downstream from the α-globin locus [58]. This element lies inside the intron of another host gene called NME4, which is up-regulated in erythroid cells through an alternative start site, as is the case for the upstream *NPRL3* gene [58,113]. Would this element take an important role when the others are missing? This is a difficult question to address, since this gene lies on another chromosome in the mouse and therefore no suitable model to test this hypothesis is currently available.

References

1. Orkin SH (1995) Hematopoiesis: how does it happen? Curr Opin Cell Biol 7: 870–877.
2. Partington KM, Jenkinson EJ, Anderson G (1999) A novel method of cell separation based on dual parameter immunomagnetic cell selection. J Immunol Methods 223: 195–205.
3. Eckert R, Mix E, von Broen B (1976) Studies on column fractionation of immune cells. VI. Separation of thymus-derived lymphocytes by means of protein beads coated with antigen-antibody complexes. Acta Biol Med Ger 35: 663–671.
4. Vernimmen D (2014) The Molecular Basis of Normal Erythroid/Megakaryocyte Development and Mechanisms of Epigenetic/Transcriptional Deregulation Leading to Erythroleukemia and Thalassaemia. In: Bonifer C, Cockerill PN, editors. Transcriptional and Epigenetic Mechanisms Regulating Normal and Aberrant Blood Cell Development Epigenetics and Human Health. Heidelberg: Springer. pp. 247–266.
5. Anguita E, Hughes J, Heyworth C, Blobel GA, Wood WG, et al. (2004) Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. EMBO J 23: 2841–2852.
6. Drissen R, Palstra RJ, Gillemans N, Splinter E, Grosveld F, et al. (2004) The active spatial organization of the beta-globin locus requires the transcription factor EKLF. Genes Dev 18: 2485–2490.
7. Vernimmen D, De Gobbi M, Sloane-Stanley JA, Wood WG, Higgs DR (2007) Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J 26: 2041–2051.
8. Fromm M, Berg P (1983) Simian virus 40 early- and late-region promoter functions are enhanced by the 72-base-pair repeat inserted at distant locations and inverted orientations. Mol Cell Biol 3: 991–999.
9. Moreau P, Hen R, Wasylyk B, Everett R, Gaub MP, et al. (1981) The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. Nucleic Acids Res 9: 6047–6068.
10. Banerji J, Rusconi S, Schaffner W (1981) Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. Cell 27: 299–308.
11. Grosschedl R, Birnstiel ML (1980) Spacer DNA sequences upstream of the T-A-T-A-A-A-T-A sequence are essential for promotion of H2A histone gene transcription in vivo. Proc Natl Acad Sci U S A 77: 7102–7106.
12. Vernimmen D, Begon D, Salvador C, Gofflot S, Grooteclaes M, et al. (2003) Identification of HTF (HER2 transcription factor) as an AP-2 (activator protein-2) transcription factor and contribution of the HTF binding site to ERBB2 gene overexpression. Biochem J 370: 323–329.
13. Hnisz D, Abraham BJ, Lee TI, Lau A, Saint-Andre V, et al. (2013) Super-enhancers in the control of cell identity and disease. Cell 155: 934–947.
14. Vo N, Goodman RH (2001) CREB-binding protein and p300 in transcriptional regulation. J Biol Chem 276: 13505–13508.
15. Carey M (1998) The enhanceosome and transcriptional synergy. Cell 92: 5–8.
16. Noonan JP, McCallion AS (2010) Genomics of long-range regulatory elements. Annu Rev Genomics Hum Genet 11: 1–23.
17. Shi J, Whyte WA, Zepeda-Mendoza CJ, Milazzo JP, Shen C, et al. (2013) Role of SWI/SNF in acute leukemia maintenance and enhancer-mediated Myc regulation. Genes Dev 27: 2648–2662.
18. Amano T, Sagai T, Tanabe H, Mizushina Y, Nakazawa H, et al. (2009) Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. Dev Cell 16: 47–57.
19. Hughes JR, Cheng JF, Ventress N, Prabhakar S, Clark K, et al. (2005) Annotation of cis-regulatory elements by identification, subclassification, and functional assessment of multispecies conserved sequences. Proc Natl Acad Sci U S A 102: 9830–9835.
20. Goh SH, Lee YT, Bhanu NV, Cam MC, Desper R, et al. (2005) A newly discovered human alpha-globin gene. Blood 106: 1466–1472.
21. Higgs DR (2013) The molecular basis of alpha-thalassemia. Cold Spring Harb Perspect Med 3: a011718.
22. Kioussis D, Vanin E, deLange T, Flavell RA, Grosveld FG (1983) Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. Nature 306: 662–666.
23. Grosveld F, van Assendelft GB, Greaves DR, Kollias G (1987) Position-independent, high-level expression of the human beta-globin gene in transgenic mice. Cell 51: 975–985.
24. Hatton CS, Wilkie AO, Drysdale HC, Wood WG, Vickers MA, et al. (1990) Alpha-thalassemia caused by a large (62 kb) deletion upstream of the human alpha globin gene cluster. Blood 76: 221–227.
25. Galanello R, Cao A (2011) Gene test review. Alpha-thalassemia. Genet Med 13: 83–88.
26. Higgs DR, Vernimmen D, Wood B (2008) Long-range regulation of alpha-globin gene expression. Adv Genet 61: 143–173.
27. Bernet A, Sabatier S, Picketts DJ, Ouazana R, Morle F, et al. (1995) Targeted inactivation of the major positive regulatory element (HS-40) of the human alpha-globin gene locus. Blood 86: 1202–1211.
28. Higgs DR, Wood WG, Jarman AP, Sharpe J, Lida J, et al. (1990) A major positive regulatory region located far upstream of the human alpha-globin gene locus. Genes Dev 4: 1588–1601.
29. Ganis JJ, Hsia N, Trompouki E, de Jong JL, DiBiase A, et al. (2012) Zebrafish globin switching occurs in two developmental stages and is controlled by the LCR. Dev Biol 366: 185–194.
30. Gillemans N, McMorrow T, Tewari R, Wai AW, Burgtorf C, et al. (2003) Functional and comparative analysis of globin loci in pufferfish and humans. Blood 101: 2842–2849.
31. Hardison R (1998) Hemoglobins from bacteria to man: evolution of different patterns of gene expression. J Exp Biol 201: 1099–1117.
32. Jarman AP, Wood WG, Sharpe JA, Gourdon G, Ayyub H, et al. (1991) Characterization of the major regulatory element upstream of the human alpha-globin gene cluster. Mol Cell Biol 11: 4679–4689.
33. Sharpe JA, Wells DJ, Whitelaw E, Vyas P, Higgs DR, et al. (1993) Analysis of the human alpha-globin gene cluster in transgenic mice. Proc Natl Acad Sci U S A 90: 11262–11266.
34. Sharpe JA, Summerhill RJ, Vyas P, Gourdon G, Higgs DR, et al. (1993) Role of upstream DNase I hypersensitive sites in the regulation of human alpha globin gene expression. Blood 82: 1666–1671.
35. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, et al. (2006) Chromatin signatures of pluripotent cell lines. Nat Cell Biol 8: 532–538.
36. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell 125: 315–326.
37. Sawarkar R, Paro R (2010) Interpretation of developmental signaling at chromatin: the Polycomb perspective. Dev Cell 19: 651–661.
38. Surface LE, Thornton SR, Boyer LA (2010) Polycomb group proteins set the stage for early lineage commitment. Cell Stem Cell 7: 288–298.
39. Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature 448: 553–560.
40. Long HK, Blackledge NP, Klose RJ (2013) ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochem Soc Trans 41: 727–740.
41. Clouaire T, Webb S, Skene P, Illingworth R, Kerr A, et al. (2012) Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. Genes Dev 26: 1714–1728.
42. Mendenhall EM, Koche RP, Truong T, Zhou VW, Issac B, et al. (2010) GC-rich sequence elements recruit PRC2 in mammalian ES cells. PLoS Genet 6: e1001244.
43. Lynch MD, Smith AJ, De Gobbi M, Flenley M, Hughes JR, et al. (2011) An interspecies analysis reveals a key role for unmethylated CpG dinucleotides in vertebrate Polycomb complex recruitment. EMBO J 31: 317–329.
44. Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, et al. (2010) CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464: 1082–1086.
45. Margueron R, Justin N, Ohno K, Sharpe ML, Son J, et al. (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. Nature 461: 762–767.
46. Wamstad JA, Alexander JM, Truty RM, Shrikumar A, Li F, et al. (2012) Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. Cell 151: 206–220.
47. Garrick D, De Gobbi M, Samara V, Rugless M, Holland M, et al. (2008) The role of the polycomb complex in silencing alpha-globin gene expression in nonerythroid cells. Blood 112: 3889–3899.
48. De Gobbi M, Garrick D, Lynch M, Vernimmen D, Hughes JR, et al. (2011) Generation of bivalent chromatin domains during cell fate decisions. Epigenetics Chromatin 4: 9.
49. Voigt P, Tee WW, Reinberg D (2013) A double take on bivalent promoters. Genes Dev 27: 1318–1338.
50. Levings PP, Zhou Z, Vieira KF, Crusselle-Davis VJ, Bungert J (2006) Recruitment of transcription complexes to the beta-globin locus control region

and transcription of hypersensitive site 3 prior to erythroid differentiation of murine embryonic stem cells. FEBS J 273: 746–755.

51. Szutorisz H, Canzonetta C, Georgiou A, Chow CM, Tora L, et al. (2005) Formation of an active tissue-specific chromatin domain initiated by epigenetic marking at the embryonic stem cell stage. Mol Cell Biol 25: 1804–1820.

52. Spooncer E, Heyworth CM, Dunn A, Dexter TM (1986) Self-renewal and differentiation of interleukin-3-dependent multipotent stem cells are modulated by stromal cells and serum factors. Differentiation 31: 111–118.

53. Grass JA, Boyer ME, Pal S, Wu J, Weiss MJ, et al. (2003) GATA-1-dependent transcriptional repression of GATA-2 via disruption of positive autoregulation and domain-wide chromatin remodeling. Proc Natl Acad Sci U S A 100: 8811–8816.

54. Marks PA, Sheffery M, Rifkind RA (1987) Induction of transformed cells to terminal differentiation and the modulation of gene expression. Cancer Res 47: 659–666.

55. Romier C, Cocchiarella F, Mantovani R, Moras D (2003) The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. J Biol Chem 278: 1336–1345.

56. Gatta R, Mantovani R (2008) NF-Y substitutes H2A-H2B on active cell-cycle promoters: recruitment of CoREST-KDM1 and fine-tuning of H3 methylations. Nucleic Acids Res 36: 6592–6607.

57. Funnell AP, Vernimmen D, Lim WF, Mak KS, Wienert B, et al. (2014) Differential regulation of the alpha-globin locus by Kruppel-like factor 3 in erythroid and non-erythroid cells. BMC Mol Biol 15: 8.

58. Lower KM, Hughes JR, De Gobbi M, Henderson S, Viprakasit V, et al. (2009) Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. Proc Natl Acad Sci U S A 106: 21771–21776.

59. Kowalczyk MS, Hughes JR, Garrick D, Lynch MD, Sharpe JA, et al. (2012) Intragenic enhancers act as alternative promoters. Mol Cell 45: 447–458.

60. Koch F, Fenouil R, Gut M, Cauchy P, Albert TK, et al. (2011) Transcription initiation platforms and GTF recruitment at tissue-specific enhancers and promoters. Nat Struct Mol Biol 18: 956–963.

61. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, et al. (2010) Widespread transcription at neuronal activity-regulated enhancers. Nature 465: 182–187.

62. De Santa F, Barozzi I, Mietton F, Ghisletti S, Polletti S, et al. (2010) A large fraction of extragenic RNA pol II transcription sites overlap enhancers. PLoS Biol 8: e1000384.

63. Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, et al. (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39: 311–318.

64. Deisseroth A, Hendrick D (1978) Human alpha-globin gene expression following chromosomal dependent gene transfer into mouse erythroleukemia cells. Cell 15: 55–63.

65. Deisseroth A, Burk R, Picciano D, Anderson WF, Nienhuis A, et al. (1975) Hemoglobin synthesis in somatic cell hybrids: globin gene expression in hybrids between mouse erythroleukemia and human marrow cells or fibroblasts. Proc Natl Acad Sci U S A 72: 1102–1106.

66. Craddock CF, Vyas P, Sharpe JA, Ayyub H, Wood WG, et al. (1995) Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. EMBO J 14: 1718–1726.

67. Johnson KD, Grass JA, Park C, Im H, Choi K, et al. (2003) Highly restricted localization of RNA polymerase II within a locus control region of a tissue-specific chromatin domain. Mol Cell Biol 23: 6484–6493.

68. Yoo EJ, Cooke NE, Liebhaber SA (2012) An RNA-independent linkage of noncoding transcription to long-range enhancer function. Mol Cell Biol 32: 2020–2029.

69. Wallace HA, Marques-Kranc F, Richardson M, Luna-Crespo F, Sharpe JA, et al. (2007) Manipulating the mouse genome to engineer precise functional syntenic replacements with human sequence. Cell 128: 197–209.

70. Devoy A, Bunton-Stasyshyn RK, Tybulewicz VL, Smith AJ, Fisher EM (2012) Genomically humanized mice: technologies and promises. Nat Rev Genet 13: 14–20.

71. Vernimmen D, Marques-Kranc F, Sharpe JA, Sloane-Stanley JA, Wood WG, et al. (2009) Chromosome looping at the human alpha-globin locus is mediated via the major upstream regulatory element (HS -40). Blood 114: 4253–4260.

72. De Gobbi M, Anguita E, Hughes J, Sloane-Stanley JA, Sharpe JA, et al. (2007) Tissue-specific histone modification and transcription factor binding in alpha globin gene expression. Blood 110: 4503–4510.

73. Vernimmen D, Lynch MD, De Gobbi M, Garrick D, Sharpe JA, et al. (2011) Polycomb eviction as a new distant enhancer function. Genes Dev 25: 1583–1588.

74. Sawado T, Halow J, Bender MA, Groudine M (2003) The beta -globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation. Genes Dev 17: 1009–1018.

75. Song SH, Kim A, Ragoczy T, Bender MA, Groudine M, et al. (2010) Multiple functions of Ldb1 required for beta-globin activation during erythroid differentiation. Blood 116: 2356–2364.

76. Ghavi-Helm Y, Klein FA, Pakozdi T, Ciglar L, Noordermeer D, et al. (2014) Enhancer loops appear stable during development and are associated with paused polymerase. Nature 512: 96–100.

77. Bender MA, Ragoczy T, Lee J, Byron R, Telling A, et al. (2012) The hypersensitive sites of the murine beta-globin locus control region act independently to affect nuclear localization and transcriptional elongation. Blood 119: 3820–3827.

78. Stadhouders R, Thongjuea S, Andrieu-Soler C, Palstra RJ, Bryne JC, et al. (2012) Dynamic long-range chromatin interactions control Myb proto-oncogene transcription during erythroid development. EMBO J 31: 986–999.

79. Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, et al. (2013) Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153: 320–334.

80. Zentner GE, Tesar PJ, Scacheri PC (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res 21: 1273–1283.

81. Ho Y, Elefant F, Liebhaber SA, Cooke NE (2006) Locus control region transcription plays an active role in long-range gene activation. Mol Cell 23: 365–375.

82. Ho Y, Elefant F, Cooke N, Liebhaber S (2002) A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. Mol Cell 9: 291–302.

83. Zhao H, Friedman RD, Fournier RE (2007) The locus control region activates serpin gene expression through recruitment of liver-specific transcription factors and RNA polymerase II. Mol Cell Biol 27: 5286–5295.

84. Spicuglia S, Kumar S, Yeh JH, Vachez E, Chasson L, et al. (2002) Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. Mol Cell 10: 1479–1487.

85. Tanimoto K, Liu Q, Bungert J, Engel JD (1999) Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. Nature 398: 344–348.

86. Blackwood EM, Kadonaga JT (1998) Going the distance: a current view of enhancer action. Science 281: 60–63.

87. Zhu X, Ling J, Zhang L, Pi W, Wu M, et al. (2007) A facilitated tracking and transcription mechanism of long-range enhancer function. Nucleic Acids Res 35: 5532–5544.

88. Hatzis P, Talianidis I (2002) Dynamics of enhancer-promoter communication during differentiation-induced gene activation. Mol Cell 10: 1467–1477.

89. Wang Q, Carroll JS, Brown M (2005) Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. Mol Cell 19: 631–642.

90. Tan M, Luo H, Lee S, Jin F, Yang JS, et al. (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell 146: 1016–1028.

91. Szutorisz H, Dillon N, Tora L (2005) The role of enhancers as centres for general transcription factor recruitment. Trends Biochem Sci 30: 593–599.

92. Stumpf M, Yue X, Schmitz S, Luche H, Reddy JK, et al. (2010) Specific erythroid-lineage defect in mice conditionally deficient for Mediator subunit Med1. Proc Natl Acad Sci U S A 107: 21541–21546.

93. Higgs DR, Vickers MA, Wilkie AO, Pretorius IM, Jarman AP, et al. (1989) A review of the molecular genetics of the human alpha-globin gene cluster. Blood 73: 1081–1104.

94. Vestri R, Pieragostini E, Ristaldi MS (1994) Expression gradient in sheep alpha alpha and alpha alpha alpha globin gene haplotypes: mRNA levels. Blood 83: 2317–2322.

95. Stamatoyannopoulos JA, Snyder M, Hardison R, Ren B, Gingeras T, et al. (2012) An encyclopedia of mouse DNA elements (Mouse ENCODE). Genome Biol 13: 418.

96. Vieira KF, Levings PP, Hill MA, Crusselle VJ, Kang SH, et al. (2004) Recruitment of transcription complexes to the beta-globin gene locus in vivo and in vitro. J Biol Chem 279: 50350–50357.

97. Ghamari A, van de Corput MP, Thongjuea S, van Cappellen WA, van Ijcken W, et al. (2013) In vivo live imaging of RNA polymerase II transcription factories in primary cells. Genes Dev 27: 767–777.

98. Sutherland H, Bickmore WA (2009) Transcription factories: gene expression in unions? Nat Rev Genet 10: 457–466.

99. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, et al. (2013) Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153: 307–319.

100. Wendt KS, Grosveld FG (2014) Transcription in the context of the 3D nucleus. Curr Opin Genet Dev 25: 62–67.

101. Illingworth RS, Bird AP (2009) CpG islands–'a rough guide'. FEBS Lett 583: 1713–1720.

102. Sen R, Grosschedl R (2010) Memories of lost enhancers. Genes Dev 24: 973–979.

103. Chen S, Ma J, Wu F, Xiong LJ, Ma H, et al. (2012) The histone H3 Lys 27 demethylase JMJD3 regulates gene expression by impacting transcriptional elongation. Genes Dev 26: 1364–1375.

104. Williams K, Christensen J, Rappsilber J, Nielsen AL, Johansen JV, et al. (2014) The Histone Lysine Demethylase JMJD3/KDM6B Is Recruited to p53 Bound Promoters and Enhancer Elements in a p53 Dependent Manner. PLoS ONE 9: e96545.

105. Taberlay PC, Kelly TK, Liu CC, You JS, De Carvalho DD, et al. (2011) Polycomb-repressed genes have permissive enhancers that initiate reprogramming. Cell 147: 1283–1294.

106. Kondo T, Isono K, Kondo K, Endo TA, Itohara S, et al. (2014) Polycomb potentiates meis2 activation in midbrain by mediating interaction of the promoter with a tissue-specific enhancer. Dev Cell 28: 94–101.

107. Seenundun S, Rampalli S, Liu QC, Aziz A, Palii C, et al. (2010) UTX mediates demethylation of H3K27me3 at muscle-specific genes during myogenesis. EMBO J 29: 1401–1411.

108. Ostuni R, Piccolo V, Barozzi I, Polletti S, Termanini A, et al. (2013) Latent enhancers activated by stimulation in differentiated cells. Cell 152: 157–171.

109. Anderson E, Peluso S, Lettice LA, Hill RE (2012) Human limb abnormalities caused by disruption of hedgehog signaling. Trends Genet 28: 364–373.

110. Sollaino MC, Paglietti ME, Loi D, Congiu R, Podda R, et al. (2010) Homozygous deletion of the major alpha-globin regulatory element (MCS-R2) responsible for a severe case of hemoglobin H disease. Blood 116: 2193–2194.

111. Coelho A, Picanco I, Seuanes F, Seixas MT, Faustino P (2010) Novel large deletions in the human alpha-globin gene cluster: Clarifying the HS-40 long-range regulatory role in the native chromosome environment. Blood Cells Mol Dis 45: 147–153.

112. Lettice LA, Daniels S, Sweeney E, Venkataraman S, Devenney PS, et al. (2011) Enhancer-adoption as a mechanism of human developmental disease. Hum Mutat 32: 1492–1499.

113. Lower KM, De Gobbi M, Hughes JR, Derry CJ, Ayyub H, et al. (2013) Analysis of sequence variation underlying tissue-specific transcription factor binding and gene expression. Hum Mutat 34: 1140–1148.