

# Upstream Distal Regulatory Elements Contact the *Lmo2* Promoter in Mouse Erythroid Cells

Anandi Bhattacharya, Chih-Yu Chen, Sara Ho, Jennifer A. Mitchell\*

Department of Cell and Systems Biology, University of Toronto, Toronto, Ontario, Canada

## Abstract

The Lim domain only 2 (*Lmo2*) gene encodes a transcriptional cofactor critical for the development of hematopoietic stem cells. Several distal regulatory elements have been identified upstream of the *Lmo2* gene in the human and mouse genomes that are capable of enhancing reporter gene expression in erythroid cells and may be responsible for the high level transcription of *Lmo2* in the erythroid lineage. In this study we investigate how these elements regulate transcription of *Lmo2* and whether or not they function cooperatively in the endogenous context. Chromosome conformation capture (3C) experiments show that chromatin-chromatin interactions exist between upstream regulatory elements and the *Lmo2* promoter in erythroid cells but that these interactions are absent from kidney where *Lmo2* is transcribed at twelve fold lower levels. Specifically, long range chromatin-chromatin interactions occur between the *Lmo2* proximal promoter and two broad regions, 3–31 and 66–105 kb upstream of *Lmo2*, which we term the proximal and distal control regions for *Lmo2* (pCR and dCR respectively). Each of these regions is bound by several transcription factors suggesting that multiple regulatory elements cooperate in regulating high level transcription of *Lmo2* in erythroid cells. Binding of CTCF and cohesin which support chromatin loops at other loci were also found within the dCR and at the *Lmo2* proximal promoter. Intergenic transcription occurs throughout the dCR in erythroid cells but not in kidney suggesting a role for these intergenic transcripts in regulating *Lmo2*, similar to the broad domain of intergenic transcription observed at the human  $\beta$ -globin locus control region. Our data supports a model in which the dCR functions through a chromatin looping mechanism to contact and enhance *Lmo2* transcription specifically in erythroid cells. Furthermore, these chromatin loops are supported by the cohesin complex recruited to both CTCF and transcription factor bound regions.

**Citation:** Bhattacharya A, Chen C-Y, Ho S, Mitchell JA (2012) Upstream Distal Regulatory Elements Contact the *Lmo2* Promoter in Mouse Erythroid Cells. PLoS ONE 7(12): e52880. doi:10.1371/journal.pone.0052880

**Editor:** Andrew C. Wilber, Southern Illinois University School of Medicine, United States of America

**Received:** August 1, 2012; **Accepted:** November 22, 2012; **Published:** December 21, 2012

**Copyright:** © 2012 Bhattacharya et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This study was supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant held by JAM). AB was funded by the Department of Cell and Systems Biology and the Department of Economics, University of Toronto. CYC was supported in part by an Ontario Graduate Scholarship. The ENCODE project is funded by the National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: ja.mitchell@utoronto.ca

## Introduction

Lim domain only 2 (LMO2) is a critical transcriptional regulator of hematopoiesis. Gene targeting experiments conducted to introduce null mutations in the mouse *Lmo2* gene, have shown that *Lmo2* is necessary for embryonic yolk sac erythropoiesis [1]. During differentiation of hematopoietic progenitor cells, *Lmo2* expression is maintained in erythroid cells but down regulated in the T-cell lineage [1,2,3,4]. Aberrant expression of *Lmo2* results in the development of T-cell related diseases; indeed *Lmo2* is located at a recurrent site of T-cell acute lymphoblastic leukemia (T-ALL) specific translocation [2,4,5,6,7,8]. In addition, patients undergoing gene therapy for X-linked severe combined immunodeficiency developed clonal T-cell proliferation as a result of aberrant transcriptional activation of *Lmo2* when the gene therapy vector integrated near *Lmo2* [9].

Previous studies have shown that in erythroid cells LMO2 is usually present as part of a complex with the transcriptional regulators, TAL1, E47, LDB1, and GATA1 [10,11,12]. This protein complex binds DNA by recognizing a bipartite DNA sequence comprising of an E box and a GATA site [10,12]. These LMO2 containing oligomeric complexes along with other factors

in hematopoietic cells have been found on the regulatory regions of various other genes including,  $\beta$ -globin (*Hbb*),  $\alpha$ -globin (*Hba*), retinaldehyde dehydrogenase 2, *c-kit* and erythroid Kruppel-like factor (*Eklf*) [13,14,15,16,17,18].

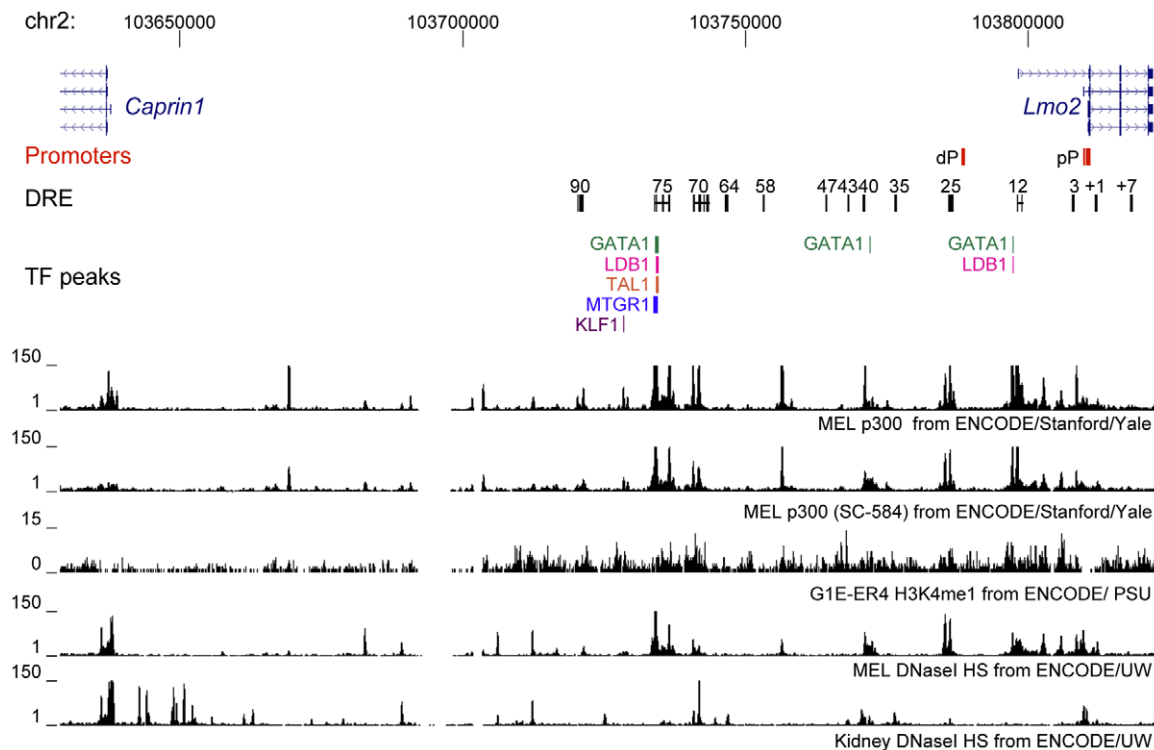
Transcription is regulated not only by the sequences immediately upstream of gene transcription start sites (TSS) but in many cases by distal regulatory elements (DRE) which can be located up or downstream of the genes they regulate. In fact genome-wide chromatin immunoprecipitation sequencing (ChIP-Seq) analysis for several transcription factors has revealed that a significant proportion (40–60%) of transcription factor bound regions are located in the intergenic regions of the genome  $\geq 10$  kb from a gene TSS [19,20,21]. In general DRE with enhancer activity are associated with increased sensitivity to DNaseI, co-binding of multiple transcription factors, binding of the histone acetyl transferase p300, increased histone H3 monomethylation of lysine 4 (H3K4me1), increased histone H3 acetylation of lysine 27 (H3K27ac) and recruitment of the mediator and cohesin complexes [21,22,23,24,25,26,27,28,29]. There are many examples of regulatory elements located several kilobases from their target genes, including the well studied regulatory elements of the *Hbb* locus control region (LCR) [30]. The LCR consists of a series

of transcription factor bound DNaseI hypersensitive sites 50 kb upstream of the *Hbb-b1* gene [31,32,33]. Chromosome conformation capture (3C) in adult erythroid cells has revealed that the *Hbb* LCR is in close proximity to the active *Hbb* genes (*Hbb-b1* and *Hbb-b2*) whereas the intervening 50 kb of DNA sequence containing the embryonic erythroid cell expressed genes is looped out [34,35,36]. The 3C technique has since been used to detect chromatin-chromatin interactions between DRE and several genes including: *Hba*, *Shh*,  $T_{H2}$ , *HoxB1* and olfactory receptor genes [37,38,39,40,41].

The chromatin-chromatin looping interactions that regulate cell-type specific gene expression are also present in a cell-type specific manner whereas many of the proteins present at sites of looping interactions are ubiquitously expressed. For example, CTCF participates in intra- and inter-chromosomal looping at individual gene loci including *Hbb*, *Igf2/H19* and *HoxA*, however, CTCF bound regions are generally bound by CTCF in all cell types [42,43,44,45]. Genome-wide studies have shown that CTCF and cohesin, a protein complex that mediates sister chromatid cohesion, localise to the same regions of the genome [46]. Furthermore, at the imprinted *IGF2-H19* locus and at the developmentally regulated *IFNG* locus both cohesin and CTCF are required for maintaining higher-order chromatin conformation [47,48]. Whereas the CTCF bound regions of the genome show limited differences between cell types, CTCF/cohesin bound regions do form tissue specific chromatin loops [42,49]. In addition, the presence of cohesin correlates with the number of

bound transcription factors at *cis*-regulatory elements not occupied by CTCF [50]. Members of the cohesin complex also interact with mediator, a complex recruited by transcription factors which acts as a bridge to the RNA polymerase II preinitiation complex [24,51]. Both mediator and cohesin proteins are bound at enhancers that form ES-cell specific chromatin loops with a nearby gene promoter [24]. In summary the cohesin complex appears to have a critical role in mediating chromatin loop formation at both CTCF bound regions of the genome and transcription factor bound distal regulatory elements required for tissue-specific gene transcription [48,52].

Transcriptional regulation of *Lmo2* in erythroid cells appears to involve multiple distal regulatory elements. Three alternate upstream promoters have been identified as well as multiple conserved DRE located up to 100 kb upstream of the *Lmo2* gene [53,54,55,56]. DRE located 90, 75, 64, 25 and 12, kb upstream of *Lmo2* and 1 kb downstream are capable of enhancing reporter gene expression in erythroid cells and in transgenic mice suggesting that strong expression of *Lmo2* in hematopoietic cells requires the combined action of upstream DRE and sequences close to the *Lmo2* proximal promoter [54]. In transgenic mice optimal expression in fetal liver and circulating erythrocytes was obtained using a multi-enhancer construct containing the 75, 70, 25, 12 and +1 DRE in conjunction with an extended proximal promoter region (pPex). Here we investigate long-range regulation of *Lmo2* transcription in the context of its endogenous genomic location using the 3C technique. Our results show that multiple



**Figure 1. Distal regulatory elements upstream of *Lmo2* overlap transcription factor bound regions in erythroid cells.** The mouse *Lmo2-Caprin1* region on chromosome 2 is depicted with chromosome coordinates shown at the top. Two *Lmo2* promoters are indicated by red boxes. Distal regulatory element (DRE) homology regions are indicated by black boxes joined by a line to delineate the human enhancer construct used in the generation of transgenic mice. Mouse ENCODE ChIP-Seq data for p300 and DNaseI hypersensitivity are shown below the DRE track. Coloured boxes represent peaks identified from transcription factor ChIP-Seq data for erythroid (MEL and G1E-ER4) cells. Overlapping transcription factor peaks were identified at the 75 and 12 DRE. These regions were also occupied by p300 and showed increased sensitivity to DNaseI. The entire locus was marked with histone H3 lysine 4 monomethylation (H3K4me1). Proximal promoter (pP), distal promoter (dP), murine erythroleukemia cells (MEL), Transcription factors (TF).  
doi:10.1371/journal.pone.0052880.g001

upstream DRE interact with the *Lmo2* proximal promoter whereas intervening regions are looped away from the proximal promoter. DRE upstream of *Lmo2* are bound by multiple transcription factors, p300, and associated with intergenic transcription when *Lmo2* is transcribed at high levels in erythroid cells. Interaction between the DRE and the *Lmo2* promoter was identified in erythroid cells but not in kidney cells suggesting a link between the looping conformation of the locus and transcriptional regulation of *Lmo2*. Furthermore, a CTCF and cohesin occupied region upstream of the most distal enhancer (90 DRE) also contacts the *Lmo2* proximal promoter region, potentially insulating the neighboring Cell cycle associated protein 1 (*Caprin1*) gene from interaction with the DRE that enhance *Lmo2* transcription.

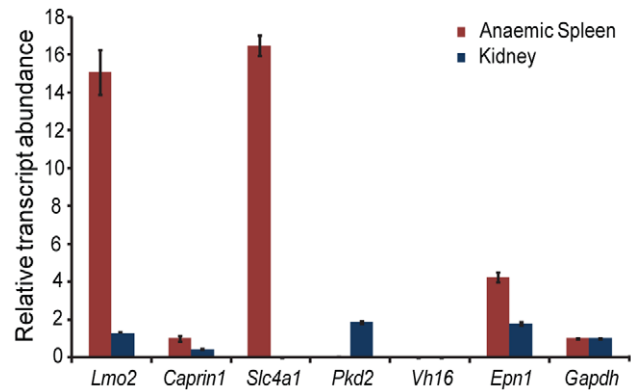
## Results

### Distal regulatory elements upstream of *Lmo2* overlap transcription factor bound regions

Multiple DRE upstream of *Lmo2* have been identified in the human genome and confirmed to have enhancer activity in transgenic mice [54]. We mapped the proximal and distal promoters and enhancer sequences identified by Landry et al. 2009 to the mouse genome and located these sequences within a region 90 kb upstream and 7 kb downstream of *Lmo2* (Figure 1, Tables S1 and S2). The intermediate promoter TSS for *Lmo2*, identified in the human genome and found to confer expression in human T-ALL cell lines that express *LMO2*, is marked by the top transcript in the gene annotation track [53]. Several transcription factors (LMO2, TAL1, GATA2, FLI1, and SFPI1) have been found associated with these DRE using a ChIP-chip approach [54]. We retrieved available ChIP-Seq data for erythroid cells (KLF1, MTGR1, GATA1, TAL1, LDB1, Table S3) to more finely map the transcription factor-bound regions within each DRE (Figure 1) [57,58,59]. Our analysis revealed several transcription factors (MTGR1, GATA1, TAL1, LDB1) bound at the 75 DRE (Figure 1). Transcription factors were also bound near the 40 (GATA1) and 12 (GATA1, LDB1) DRE. In addition to transcription factors, DRE with enhancer activity are associated with p300 and increased DNaseI sensitivity [22,25,26,60]. Using available ChIP-Seq data from the mouse ENCODE project (Table S3) [61] we identified p300 and DNaseI hypersensitivity specific to erythroid cells at the 75, 25 and 12 DRE. In contrast H3K4me1, another feature associated with enhancer regions, was found to be generally increased through the 100 kb region encompassing the previously identified DREs near *Lmo2*.

### The 75 distal regulatory element contacts the *Lmo2* proximal promoter

To investigate whether or not chromatin loops form in erythroid cells which bring the 75 DRE into proximity with the *Lmo2* promoter we performed 3C in adult erythroid cells isolated from mouse anemic spleens 5 days after the initiation of Phenylhydrazine treatment [62]. On day five the anaemic spleen is composed of >85% globin expressing erythroid cells [63]. For comparison we used kidney as a tissue in which *Lmo2* is not transcribed at robust levels. We confirmed robust transcription of *Lmo2* in isolated anaemic spleen by measuring the levels of the primary transcript by RT-qPCR (Figure 2). By contrast *Lmo2* primary transcript levels were twelve fold lower in kidney. We also examined the primary transcript levels of the cell cycle associated protein *Caprin1* that is located 172 Kb upstream of *Lmo2* with the DREs upstream of *Lmo2* located between the two genes. We found that *Caprin1* is transcribed in both adult erythroid and kidney cells



**Figure 2. *Lmo2* primary transcripts are abundant in erythroid cells.** Primary transcript levels in adult mouse anaemic spleen (red) and kidney (blue) for: *Lmo2* (exon2-intron2), *Caprin1* (exon3-intron2), *Slc4a1* (exon1-intron1), *Pkd2* (intron2-exon3), *Epn1* (exon1-intron1), *Gapdh* (exon1-intron1) and *Vh16* (*genic*). Levels were quantitatively assessed by RT-qPCR and expressed relative to *Gapdh*. *Epn1* is a second ubiquitously expressed reference gene, *Slc4a1* is an erythroid cell specific transcript, *Pkd2* is a kidney specific transcript, *Vh16* is not expressed in either tissue.

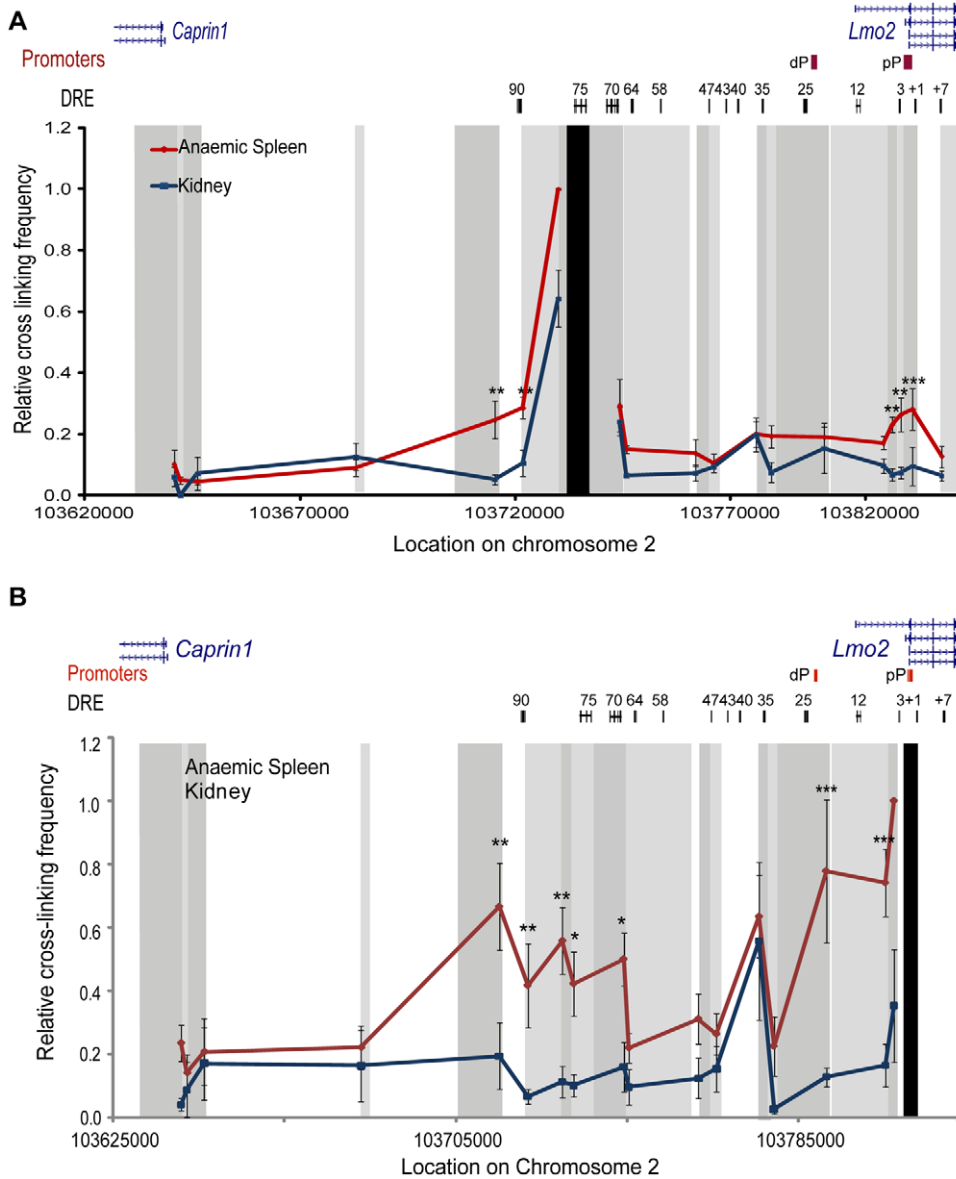
doi:10.1371/journal.pone.0052880.g002

with primary transcript levels 2 fold higher in adult erythroid cells compared to kidney.

As our analysis of the ChIP-Seq data revealed the highest density of transcription factors bound at the 75 DRE we performed 3C experiments using the *HindIII* fragment containing the 75 DRE as the anchor fragment (Figure 3A, restriction fragment map and primers shown in Figure S1). 3C analysis revealed significantly increased interaction between the 75 DRE and the *Lmo2* proximal promoter fragment in erythroid cells compared to kidney tissues. Of note the *Lmo2* proximal promoter fragment also contains the +1 enhancer element found to cooperate with the 75 DRE for optimal expression in circulating erythrocytes of transgenic mice [54]. We also found increased interaction of the 75 DRE with the two fragments upstream of the *Lmo2* proximal promoter in erythroid cells compared to kidney. This region contains the 3 DRE identified by Landry et al. 2009 which appeared not to have enhancer activity in transgenic analysis. We did not identify interaction of the 75 DRE with the fragment containing the 25 DRE/distal promoter or the 12 DRE/intergenic promoter. Interestingly, the 75 DRE was found to have significantly increased interaction with a fragment overlapping the 90 DRE and a fragment upstream of the 90 DRE in erythroid cells compared to kidney. We found no interaction between the 75 DRE and the *Caprin1* TSS fragment. These results indicate that chromatin-chromatin interactions between 75 DRE and the *Lmo2* proximal promoter and +1 enhancer occur when *Lmo2* is transcribed in erythroid cells but not in kidney where the gene is transcribed at twelve fold lower levels (Figure 2).

### Several upstream distal regulatory elements contact the *Lmo2* promoter

Next we performed a locus-wide 3C using the *Lmo2* proximal promoter as the anchor fragment. We found significantly increased interaction between the *Lmo2* proximal promoter and the fragments containing the 12, 25, 70, 75 and 90 DRE in erythroid cells compared to kidney cells (Figure 3B). These findings indicate that two broad domains interact with the *Lmo2* proximal promoter in a tissue specific manner. The first domain, which we term the proximal control region (pCR), is located close to the *Lmo2*



**Figure 3. Distal regulatory elements interact with the *Lmo2* proximal promoter in erythroid cells.** A) Quantitative chromosome conformation capture (3C) was performed to detect chromatin-chromatin interactions between the 75 DRE (distal regulatory element) upstream of *Lmo2* and the rest of the *Lmo2-Caprin1* region of mouse chromosome 2. B) Similarly, 3C was performed to detect chromatin-chromatin interactions between the *Lmo2* proximal promoter (pP) and distal regulatory elements (DRE). In both the profile of interactions identified in anaemic spleen (red) and kidney (blue) is displayed. Black box indicates the anchor fragment at the 75 DRE or the *Lmo2* pP and alternating intensities of grey boxes indicate the fragments investigated for interactions. Data points are an average of three to five independent biological replicates. Error bars depict the SEM, \*  $p < 0.05$ , \*\*  $p < 0.01$ , and \*\*\*  $p < 0.001$ . doi:10.1371/journal.pone.0052880.g003

proximal promoter and contains the distal and intermediate promoters as well as the 12 and 25 DRE. The second interaction domain located further upstream of *Lmo2*, which we term the distal control region (dCR), contains the 70, 75 and 90 DRE as well as a fragment upstream of the 90 DRE. These are similar, though not identical, to clusters I (-90 to -64) and II (-40 to +1) identified by Landry et al. 2009 as enriched in histone H3 acetylation in erythroid cells [54]. We also detected a peak in the relative interaction frequency at the 35 DRE; however this interaction was detected in both kidney and erythroid cells (Figure 3B).

Our initial analysis of transcription factors bound in the intergenic region between *Caprin1* and *Lmo2* highlighted the 75

DRE as being bound by multiple transcription factors but did not reveal binding at the 90 and 70 DRE. By contrast our 3C data revealed a broad domain of chromatin-chromatin contacts between the 70, 75 and 90 DRE and the *Lmo2* proximal promoter raising questions about the function of this dCR. We did identify additional p300 association at the 70 DRE suggesting that bound transcription factors are present which recruit p300 to this DRE (Figure 1). To investigate additional transcription factors bound upstream of *Lmo2* we retrieved transcription factor ChIP-Seq data for the HPC7 hematopoietic progenitor cell line: ERG, FLI1, GATA2, GFI1B, LMO2, MEIS1, PU1, TAL1, and RUNX1 [64]. This data revealed multiple transcription factors associated with

other DRE in the intergenic region including the 90, 75, 70, 64, 40, 25 and 12 DRE (Figure S2).

### The *Caprin1* TSS does not interact with the identified distal regulatory elements

*Caprin1* is a ubiquitously expressed gene located 172 kb upstream of *Lmo2* and transcribed from the opposite strand. As *Caprin1* is transcribed in erythroid cells we were interested to investigate whether or not *Caprin1* physically interacts with the DRE located between *Caprin1* and *Lmo2*. Our initial 3C experiments performed with the 75 DRE as the anchor fragment did not show any increase in the relative interaction frequency between the *Caprin1* TSS and the 75 DRE (Figure 3A), however as *Caprin1* could physically interact with other DRE located between *Caprin1* and *Lmo2* we performed 3C experiments using the *HindIII* fragment containing the *Caprin1* TSS as the anchor fragment. Our results did not show any significant peaks in the relative interaction frequency with the *Caprin1* TSS in the entire region between the *Caprin1* and *Lmo2* genes (Figure S3). Furthermore we identified no significant differences between the relative interaction frequency of any *HindIII* fragments with the *Caprin1* TSS for cells isolated from anaemic spleen and kidney.

### CTCF and Rad21 are bound within regions of chromatin-chromatin interaction

Binding of CTCF and recruitment of the cohesin complex appears to be involved in many instances of chromatin-chromatin interactions at other loci [42,43,44,45,46,47,48]. Investigating ChIP-Seq data released by the mouse ENCODE project (Table S3) [61] we identified several CTCF and cohesin (RAD21) bound regions within the dCR (Figure 4) specifically within the fragment upstream of the 90 DRE. However, unlike the chromatin looping interactions, which we identified as being specific to erythroid cells, the CTCF bound region upstream of the 90 DRE was bound by CTCF in several cell types (Figure S4). Additional CTCF and cohesin bound regions are located just down-stream of the 75 DRE and at the *Lmo2* proximal promoter. Like the CTCF bound region upstream of the 90 DRE, CTCF was bound at the *Lmo2* proximal promoter in several cell types. CTCF bound 400 bp downstream of the 75 DRE was most

prominent in erythroid cells but also observed in bone marrow, heart and lung tissues (Figure S4).

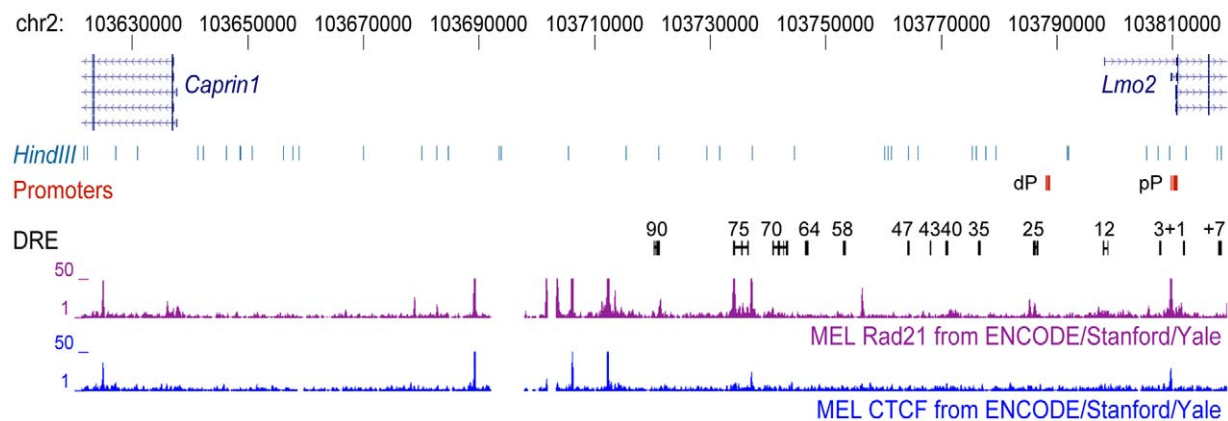
### Intergenic transcription at the distal regulatory elements

Previous studies have identified intergenic transcripts throughout the human Hbb LCR which functions via a chromatin looping mechanism to regulate the  $\beta$ -globin genes [65,66,67]. In addition enhancer RNA (eRNA) has been identified at several neuronal enhancers [68]. Analysis of RNA-Seq data from mouse fetal liver erythroblasts [69] revealed intergenic transcription upstream of *Lmo2*, between the 58 and 70 DREs (Figure 5). To investigate intergenic transcription within this region in adult anaemic spleen erythroid cells and in kidney tissue we performed RT-qPCR. This analysis identified measurable levels of intergenic transcription occurring at all identified DRE in erythroid cells (Figure 5, Table S4). By contrast little or no intergenic transcription was detected in kidney tissue. In erythroid cells we detected the highest levels of intergenic transcription at the 58 DRE which diminished towards the 90 DRE, covering most of the dCR, whereas the intervening regions between the 58 and 12 DRE showed considerably lower levels of intergenic transcription.

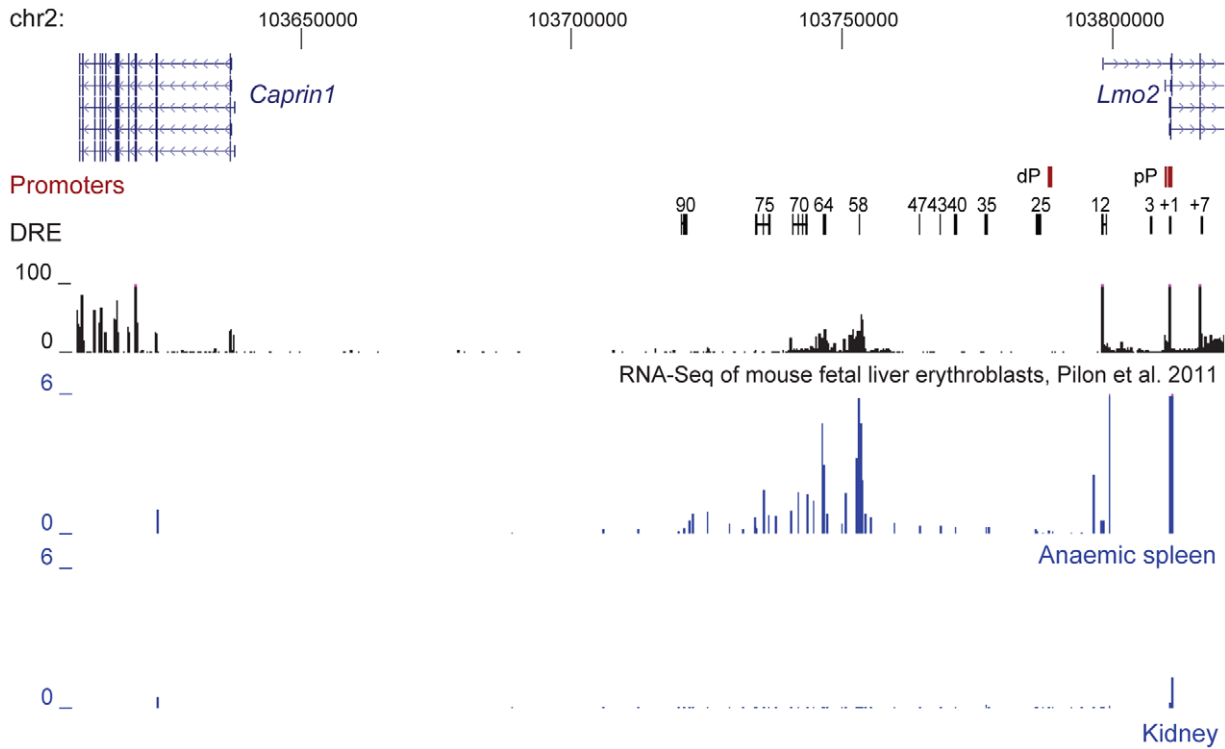
### Discussion

Our investigation of chromatin-chromatin interactions throughout the mouse genomic region containing *Lmo2* and *Caprin1* identified several erythroid cell specific interactions between the *Lmo2* proximal promoter and upstream DRE. Specifically we identified chromatin-chromatin interactions in erythroid cells between the *Lmo2* proximal promoter and both a distal interaction cluster (dCR), containing three transcription factor bound DRE (70, 75, 90), and a proximal interaction cluster (pCR). Furthermore, 3C experiments revealed no significant interactions between the *Caprin1* TSS and the DRE suggesting these elements are specific in regulating *Lmo2* transcription.

Using a combination of ChIP-Seq data for mature erythroid cells and HPC7 hematopoietic progenitor cells we identified multiple transcription factors bound within both the pCR and the dCR. Even considering the HPC7 ChIP-Seq data that showed more transcription factor peaks across the entire region upstream of *Lmo2* the 75 DRE contained the highest density of transcription factor association as well as p300, RAD21 and a nearby CTCF



**Figure 4. CTCF and RAD21 are bound within the *Lmo2-Caprin1* region at sites of chromatin-chromatin interaction.** The mouse *Lmo2-Caprin1* region on chromosome 2 is depicted with chromosome coordinates shown at the top. *Lmo2* promoters are indicated by red boxes. Distal regulatory element (DRE) homology regions are indicated by black boxes joined by a line to delineate the human enhancer construct used in the generation of transgenic mice. Mouse ENCODE ChIP-Seq data for the cohesin complex member RAD21 and CTCF are shown below DRE. Proximal promoter (pP), distal promoter (dP), murine erythroleukemia cells (MEL differentiated with 2% DMSO). doi:10.1371/journal.pone.0052880.g004



**Figure 5. Intergenic transcription occurs upstream of *Lmo2* in erythroid cells.** The mouse *Lmo2-Caprin1* region on chromosome 2 is depicted with chromosome coordinates shown at the top. RNA-Seq data for mouse fetal liver erythroblasts from Pilon et al. 2011 was obtained from the PSU Genome Browser (replicate 1 is shown in black). Transcript levels in adult mouse anaemic spleen and kidney were quantitatively assessed by RT-qPCR (shown in blue) and depicted relative to *Gapdh*. The levels downstream of 12 DRE (21.57) and the *Lmo2* pP (15.07) relative to *Gapdh* are off scale. Distal regulatory elements (DRE), distal promoter (dP), proximal promoter (pP). doi:10.1371/journal.pone.0052880.g005

bound region. Furthermore our 3C data confirmed specific interaction of the 75 DRE with the *Lmo2* proximal promoter and interaction of the *Lmo2* proximal promoter with the dCR containing the 75 DRE as well as the 90 and 70 DRE. In addition, a CTCF bound region upstream of the 90 DRE is contained in the dCR. Previous studies in circulating erythroid cells of transgenic mice show that the 75 DRE has the strongest enhancer activity in erythroid cells and drives the expression of a reporter gene, cooperatively with the *Lmo2* proximal promoter and +1 enhancer element [54]. We have shown that the 75 DRE functions via contacting the *Lmo2* proximal promoter/+1 DRE region to form an erythroid cell specific chromatin loop that includes other regulatory elements and a cluster of CTCF bound regions. This suggests that in the endogenous context elements throughout the dCR coordinately regulate *Lmo2* transcription in erythroid cells.

We also identified increased interaction throughout a pCR containing the 25 DRE/distal promoter and 12 DRE/intergenic promoter and the +1DRE/proximal promoter in erythroid cells compared to kidney cells, suggesting that this more proximal region also contributes to transcriptional regulation of endogenous *Lmo2*. Interestingly, we did not identify interactions of the 75 DRE with the fragments containing the 25 or 12 DRE. Transgenic analysis by Landry et al. 2009 of these elements did reveal that the 25 and 12 DRE conferred expression in the fetal liver whereas only the 75 DRE conferred expression in fetal liver as well as in circulating blood cells suggesting that these elements have different functional roles in regulating *Lmo2* expression. The fact that we did not identify specific interactions between the 75 DRE and the fragment containing the 25 or 12 DRE suggests that the interactions of these two regions with the proximal promoter are

mutually exclusive. These mutually exclusive interactions could occur in different sub-populations of cells within the anaemic spleen or the interactions could be dynamic within individual cells with the proximal promoter region alternately contacting the 25–12 region and the 70–90 region similar to the flip-flop of the *Hbb* LCR between the  $\gamma$ - and  $\beta$ -genes [70].

We identified erythroid-cell specific chromatin loops between a dCR and the *Lmo2* proximal promoter; however the question remains as to which factors are mediating these looping interactions. LMO2 itself is one of the regulatory transcription factors bound to the upstream DRE, specifically at 75, 25 and 12, all of which showed increased interaction with the *Lmo2* promoter suggesting an important role for the LMO2 complex in chromatin looping. In support of this, a recent study found LDB1 (a member of the LMO2 complex) at regions of chromatin interaction with the LDB1 bound *Hbb-b1* promoter [57]. GATA1 and KLF1 are also bound within the distal interacting region, these transcription factors have been shown to be required, though not sufficient for chromatin looping between *Hbb-b1* and the LCR [44,71,72] and may have a similar role in regulating looping within the *Lmo2* locus. Cohesin (RAD21) is bound within both the dCR and pCR, specifically at the 75, 25 DRE as well as at the proximal promoter suggesting cohesin bound at the upstream DRE supports erythroid cell specific looping interactions with the *Lmo2* proximal promoter.

Cohesin is also recruited to CTCF occupied regions [46] and we identified several CTCF/RAD21 bound regions throughout the *Lmo2* upstream region, all of which occur within erythroid cell specific interacting domains, though the majority of the CTCF sites were not specific to erythroid cells. This is similar to the findings at the *Hbb* locus where CTCF bound regions, invariant

between cell types, formed cell type specific chromatin loops [49]. CTCF is often associated with insulator activity; the CTCF/RAD21 occupied region upstream of the 90 DRE could have an important role in preventing *Caprin1* from interacting with the DRE that enhance *Lmo2* transcription in erythroid cells. As overexpression of *Caprin1* causes inhibition of cell division it is critical to prevent its aberrant upregulation in rapidly cycling erythroid cells [73]. In support of an important role for the CTCF bound region upstream of the 90 DRE this region is conserved and bound by CTCF in the human genome (not shown). We did identify one CTCF/RAD21 bound region, downstream of the 75 DRE, which may be critical in generating the tissue specific looping pattern that we identified. In summary, our results are consistent with chromatin-chromatin interactions throughout the *Lmo2* locus being supported by cohesin recruited both to CTCF bound regions (upstream of 90, downstream of 75) as well as at transcription factor and p300 bound enhancers not associated with CTCF (90, 70, 25, and 12 DRE).

Our analysis indicates that a large portion of the dCR is transcribed at moderate levels in erythroid cells but not in kidney. Whereas short eRNAs have been identified at enhancers in neuronal cells the broad domain of intergenic transcription we observed throughout the *Lmo2* dCR is more reminiscent of the human *Hbb* LCR [65,66,67,68]. Intergenic transcription throughout the *Lmo2* dCR may function to facilitate the physical interaction of the dCR with the *Lmo2* proximal promoter through recruitment of both regions to a shared transcription factory. This co-localization at a shared transcription compartment would then allow transcription factors recruited to the dCR to influence the basal transcriptional machinery recruited to the *Lmo2* proximal promoter thereby enhancing *Lmo2* transcription.

In conclusion, we found that the mouse *Lmo2-Caprin1* locus adopts a tissue-specific conformation in erythroid cells. This tissue specific organization of the locus brings several DRE into proximity with the *Lmo2* proximal promoter while excluding the *Caprin1* TSS. A proximal control region, immediately upstream of the *Lmo2* proximal promoter, contains the 12 and 25 DRE as well as the distal and intermediate promoters and is more closely associated with the *Lmo2* proximal promoter in erythroid cells compared to kidney. Furthermore, a distal control region, covering 39 kb and containing three DRE (90, 75 and 70), forms a strong interaction with the *Lmo2* proximal promoter and shares many features with the well characterized *Hbb* LCR. A CTCF bound region upstream of the 90 DRE flanks the dCR and may function as an insulator preventing the interaction of *Caprin1* with the erythroid cell specific *Lmo2* enhancers.

## Materials and Methods

### Ethics Statement

All experiments were approved by the University Animal Care Committee (UACC) at the University of Toronto and the Bioscience Local Animal Care Committee (LACC).

### Cell Isolation

Adult globin expressing mature erythroid cells were isolated from C57/Blk6 mice in large numbers ( $>1 \times 10^6$ ) from the spleen of mice 5 days after phenylhydrazine treatment. This treatment induces haemolytic anaemia, as a result of which the spleen becomes the major site of red blood cell production [62]. We disrupted fresh spleen or kidney tissue into a single-cell suspension and processed cells immediately as detailed below.

## Chromosome Conformation Capture (3C)

3C experiments were performed as described by Dekker et al. 2002 with the following modifications [74]. Formaldehyde was added to 2%, and the samples were cross linked for 10 minutes at room temperature. An overnight ligation of the digested DNA with T4 DNA ligase was performed. The 3C control template was prepared by mixing equimolar amounts of the BAC clone of the entire *Lmo2-Caprin1* locus (RP23-76D2) with the *Alpha Aortic Actin 2* BAC clone (RP23-2N15) followed by digestion with *HindIII*. The digested DNA was then ligated, and purified using phenol/chloroform extraction and ethanol precipitation. *HindIII* restriction enzyme digestion efficiency was confirmed to be between 85 and 95% efficient at several genomic fragments in anaemic spleen and kidney cells (Figure S5).

The linear range of amplification was determined for erythroid and kidney samples by serial dilution. An appropriate amount of the DNA within the linear range (typically 40 ng of DNA) was subsequently used for quantification. PCR products of the ligated fragments were quantified using real-time quantitative PCR (qPCR) on the Bio-Rad CFX-384 cyclor. All data points were generated from an average of three-five independent 3C experiments with the qPCR performed in triplicate. Standard curves were generated by 5 fold serial dilution of the 3C control template and run in parallel with 3C experimental samples. The primers used in qPCR are listed in Table S5. In each individual experiment 3C data were normalized to neighbouring fragments at the Alpha aortic actin locus ( $\alpha$ -A2).

## RNA isolation and real-time RT-qPCR

RNA from anaemic spleen and kidney was isolated using TRIzol, according to the manufacturer's instructions (Invitrogen). The iScript First strand synthesis cDNA kit from Bio-Rad was used for preparation of random-hexamer primed cDNA. Amplification in qPCR was performed on the Bio-Rad CFX-384 cyclor using the standard curve method. The reaction mixture contained 2X Bio-Rad iTaq SYBR green mastermix, 0.3 pM of each primer, 1uL cDNA (10 times diluted from a 20 uL of reverse transcription reaction). The conditions for qPCR were as follows: 94°C for 3 min followed by 40 cycles at 94°C for 30 s, 62°C for 30 s. Expression levels of *Gapdh* or *Epn1* were used for normalization of expression levels. The primers used for real-time RT-qPCR are listed in Table S5.

## Statistical analysis

The 3C data were analyzed by two-way ANOVA using Sigma Plot12. Post tests (Holm-Sidak method) were performed to assess significant differences between anaemic spleen and kidney samples at specific genomic locations.

## Genome Mapping and Peak Identification from ChIP-Seq datasets in erythroid cells

ChIP-Seq raw data for GATA1, KLF1, LDB1, TAL1, and MTGR1 [57,58,59] listed in Table S3 were downloaded from Gene Expression Omnibus (GEO) [75]. ChIP-Seq data were aligned to NCBI mm37 mouse assembly (mm9) using Bowtie alignment [76] by suppressing alignments to only 1 best reportable alignment with a maximum number of 2 mismatches within 28 nucleotides of seed length in the high quality end. The SISSRs [77] algorithm was subsequently used to identify significant transcription factor peaks compared to that of the input DNA with  $p < 0.001$ . To remove amplification bias, multiple reads aligning to the same genomic coordinate were counted as one. Parameters for the corresponding transcription factor data were

set according to original publications using applicable input data sets. Significant transcription factor peaks were uploaded to UCSC genome browser for visualization [78]. The HPC7 ChIP-Seq data analysis was performed by using published peaks [64]. ChIP-Seq data for CTCF, p300, RAD 21, H3K4me1, DNaseI hypersensitivity data were obtained from the mouse ENCODE project (Table S3) [61].

## Supporting Information

**Figure S1 The *Lmo2/Caprin1* region on mouse chromosome 2.** Primers used in chromosome conformation capture (3C) and *HindIII* restriction sites are shown across the *Lmo2/Caprin1* region of mouse chromosome 2. Promoters and distal regulatory elements (DRE) are depicted in red and black respectively. Anchor fragments used in the *Caprin1*, 75 DRE and *Lmo2* 3C experiments are marked with an asterisk (\*). Distal promoter (pP), proximal promoter (pP). (PDF)

**Figure S2 Distal regulatory elements upstream of *Lmo2* overlap transcription factor bound regions in HPC7 hematopoietic progenitor cells.** The mouse *Lmo2-Caprin1* region. Distal regulatory element (DRE) homology regions are indicated by black boxes joined by a line to delineate the human enhancer construct used in the generation of transgenic mice. Coloured boxes represent peaks identified from transcription factor ChIP-Seq data from HPC7 hematopoietic progenitor cells obtained from Wilson et al. 2010. Proximal promoter (pP), distal promoter (dP). (PDF)

**Figure S3 The *Caprin1* TSS does not interact with distal regulatory elements upstream of *Lmo2*.** Quantitative chromosome conformation capture (3C) was performed to detect chromatin-chromatin interactions between the *Caprin1* TSS and distal regulatory elements (DRE). The profile of interactions identified in anaemic spleen (red) and kidney (blue) is displayed. Black box indicates the anchor fragment at *Caprin1* and alternating intensities of grey boxes indicate the fragments investigated for interactions. Data points are an average of three independent biological replicates. Error bars depict the SEM, no significant differences were identified throughout this region. (PDF)

**Figure S4 CTCF bound upstream of *Lmo2* in different cell types.** The mouse *Lmo2* upstream region on chromosome 2 is depicted with chromosome coordinates shown at the top. *HindIII* restriction sites are indicated by blue lines. The two *Lmo2* promoters are indicated by red boxes. Distal regulatory element (DRE) homology regions are indicated by black boxes joined by a line to delineate the human enhancer construct used in the generation of transgenic mice. Mouse ENCODE ChIP-Seq data from B Ren (Ludwig Inst. for Cancer Research) and M Snyder (Stanford University) for CTCF in different cell types are shown below the DRE. Proximal promoter (pP), distal promoter (dP), murine erythroleukemia cells (MEL differentiated with 2% DMSO), bone marrow (BM), embryonic stem cells (ES-Bruce4), mouse embryonic fibroblasts (MEF). (PDF)

## References

- Warren AJ, Colledge WH, Carlton MB, Evans MJ, Smith AJ, et al. (1994) The oncogenic cysteine-rich LIM domain protein rbtn2 is essential for erythroid development. *Cell* 78: 45–57.
- Boehm T, Foroni L, Kaneko Y, Perutz MF, Rabbitts TH (1991) The rhombotin family of cysteine-rich LIM-domain oncogenes: distinct members are involved in T-cell translocations to human chromosomes 1p15 and 11p13. *Proc Natl Acad Sci U S A* 88: 4367–4371.

**Figure S5 Restriction digestion efficiency in chromosome conformation capture.** Restriction digestion efficiency was between 85 and 95% at several *HindIII* restriction sites. *Lmo2* proximal promoter (pP), Distal regulatory element (DRE), *Caprin1* promoter (*Caprin1P*). “U” denotes a restriction fragment upstream of the indicated element.

(PDF)

**Table S1 Coordinates of distal regulatory elements located upstream of the *Lmo2* promoter in the mouse genome.** Distal regulatory elements are named according to their distance upstream of the annotated *Lmo2* transcription start site overlapping the proximal promoter. Coordinates are given for homology regions identified by BLAT. All fragments were mapped in NCBI m37 mouse assembly (mm9).

(PDF)

**Table S2 Coordinates of the *Lmo2* proximal and distal promoters in the mouse genome.** The coordinates of proximal promoters and distal promoters for the *Lmo2* gene in the mouse genome are listed in the table. Coordinates are given for homology regions identified by BLAT. All fragments were mapped in NCBI m37 mouse assembly (mm9).

(PDF)

**Table S3 Chromatin immunoprecipitation sequencing data.** Transcription factor binding sites have been obtained from three different cell types; differentiated murine erythroleukemia cells (MEL), hematopoietic progenitor cells (HPC7), and GIE-ER4 a GATA1-null erythroblast cell line in which GATA1 activity was restored. CTCF, DNaseI hypersensitivity, p300 and RAD21 data have been obtained from the mouse ENCODE project, sources listed (Principal investigator, Institution).

(PDF)

**Table S4 Relative transcript abundance in intergenic regions upstream of *Lmo2*.**

(PDF)

**Table S5 Primers.** Specific primers are listed for the chromosome conformation capture (3C) and RT-qPCR analyses. Left primer (L), right primer (R), primers used to test *HindIII* restriction digestion efficiency are marked as REX.

(PDF)

## Acknowledgments

We would like to thank all the members of the Mitchell lab for helpful discussions specifically Dr. Neil Macpherson, Dr. Scott Davidson and Michael Schwartz for critical review of the manuscript. We would also like to thank the members of the ENCODE project for generating and releasing data to the scientific community. Specifically we acknowledge the contributions of Drs B Ren (Ludwig Inst. for Cancer Research), M Snyder (Stanford University), R Hardison (Penn State University) and JA Stamatoyannopoulos (University of Washington) for generating genome-wide data used in this study.

## Author Contributions

Conceived and designed the experiments: JAM AB. Performed the experiments: AB SH. Analyzed the data: AB CYC JAM. Wrote the paper: AB JAM.



3. Foroni L, Boehm T, White L, Forster A, Sherrington P, et al. (1992) The rhombotin gene family encode related LIM-domain proteins whose differing expression suggests multiple roles in mouse development. *J Mol Biol* 226: 747–761.
4. Royer-Pokora B, Loos U, Ludwig WD (1991) TTG-2, a new gene encoding a cysteine-rich protein with the LIM motif, is overexpressed in acute T-cell leukaemia with the t(11;14)(p13;q11). *Oncogene* 6: 1887–1893.
5. Fisch P, Boehm T, Lavenir I, Larson T, Arno J, et al. (1992) T-cell acute lymphoblastic lymphoma induced in transgenic mice by the RBTN1 and RBTN2 LIM-domain genes. *Oncogene* 7: 2389–2397.
6. Fitzgerald TJ, Neale GA, Raimondi SC, Goorha RM (1992) Rhom-2 expression does not always correlate with abnormalities on chromosome 11 at band p13 in T-cell acute lymphoblastic leukemia. *Blood* 80: 3189–3197.
7. Larson RC, Fisch P, Larson TA, Lavenir I, Langford T, et al. (1994) T cell tumours of disparate phenotype in mice transgenic for Rbm-2. *Oncogene* 9: 3675–3681.
8. Garcia IS, Kaneko Y, Gonzalez-Sarmiento R, Campbell K, White L, et al. (1991) A study of chromosome 11p13 translocations involving TCR beta and TCR delta in human T cell leukaemia. *Oncogene* 6: 577–582.
9. Haccin-Bey-Abina S, Von Kalle C, Schmidt M, McCormack MP, Wulffraat N, et al. (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* 302: 415–419.
10. Osada H, Grutz G, Axelson H, Forster A, Rabbitts TH (1995) Association of erythroid transcription factors: complexes involving the LIM protein RBTN2 and the zinc-finger protein GATA1. *Proc Natl Acad Sci U S A* 92: 9585–9589.
11. Valge-Archer VE, Osada H, Warren AJ, Forster A, Li J, et al. (1994) The LIM protein RBTN2 and the basic helix-loop-helix protein TAL1 are present in a complex in erythroid cells. *Proc Natl Acad Sci U S A* 91: 8617–8621.
12. Wadman IA, Osada H, Grutz GG, Agulnick AD, Westphal H, et al. (1997) The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J* 16: 3145–3157.
13. Ono Y, Fukuhara N, Yoshie O (1998) TAL1 and LIM-only proteins synergistically induce retinaldehyde dehydrogenase 2 expression in T-cell acute lymphoblastic leukemia by acting as cofactors for GATA3. *Mol Cell Biol* 18: 6939–6950.
14. Song SH, Hou C, Dean A (2007) A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Mol Cell* 28: 810–822.
15. Song SH, Kim A, Ragozy T, Bender MA, Groudine M, et al. (2010) Multiple functions of Ldb1 required for beta-globin activation during erythroid differentiation. *Blood* 116: 2356–2364.
16. Anderson KP, Crable SC, Lingrel JB (1998) Multiple proteins binding to a GATA-E box-GATA motif regulate the erythroid Kruppel-like factor (EKLF) gene. *J Biol Chem* 273: 14347–14354.
17. 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.
18. Lecuyer E, Herblot S, Saint-Denis M, Martin R, Begley CG, et al. (2002) The SCL complex regulates c-kit expression in hematopoietic cells through functional interaction with Sp1. *Blood* 100: 2430–2440.
19. Yu M, Riva L, Xie H, Schindler Y, Moran TB, et al. (2009) Insights into GATA-1-mediated gene activation versus repression via genome-wide chromatin occupancy analysis. *Mol Cell* 36: 682–695.
20. Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, et al. (2009) An oestrogen-receptor-alpha-bound human chromatin interactome. *Nature* 462: 58–64.
21. Chen X, Xu H, Yuan P, Fang F, Huss M, et al. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133: 1106–1117.
22. Gross DS, Garrard WT (1988) Nuclease hypersensitive sites in chromatin. *Annu Rev Biochem* 57: 159–197.
23. Chen CY, Morris Q, Mitchell JA (2012) Enhancer identification in mouse embryonic stem cells using integrative modeling of chromatin and genomic features. *BMC Genomics* 13: 152.
24. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, et al. (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467: 430–435.
25. Visel A, Blow MJ, Li Z, Zhang T, Akiyama JA, et al. (2009) ChIP-seq accurately predicts tissue-specific activity of enhancers. *Nature* 457: 854–858.
26. Blow MJ, McCulley DJ, Li Z, Zhang T, Akiyama JA, et al. (2010) ChIP-Seq identification of weakly conserved heart enhancers. *Nat Genet* 42: 806–810.
27. Creighton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, et al. (2010) Histone H3K27ac separates active from poised enhancers and predicts developmental state. *Proc Natl Acad Sci U S A* 107: 21931–21936.
28. Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, et al. (2011) A unique chromatin signature uncovers early developmental enhancers in humans. *Nature* 470: 279–283.
29. 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.
30. Tuan DY, Solomon WB, London IM, Lee DP (1989) An erythroid-specific, developmental-stage-independent enhancer far upstream of the human “beta-like globin” genes. *Proc Natl Acad Sci U S A* 86: 2554–2558.
31. Forrester WC, Thompson C, Elder JT, Groudine M (1986) A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc Natl Acad Sci U S A* 83: 1359–1363.
32. Moon AM, Ley TJ (1990) 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* 87: 7693–7697.
33. Tuan D, Solomon W, Li Q, London IM (1985) The “beta-like-globin” gene domain in human erythroid cells. *Proc Natl Acad Sci U S A* 82: 6384–6388.
34. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P (2002) Long-range chromatin regulatory interactions in vivo. *Nat Genet* 32: 623–626.
35. Palstra RJ, Tolhuis B, Splinter E, Nijmeijer R, Grosveld F, et al. (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet* 35: 190–194.
36. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10: 1453–1465.
37. Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, et al. (2006) Interchromosomal interactions and olfactory receptor choice. *Cell* 126: 403–413.
38. Spilianakis CG, Flavell RA (2004) Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5: 1017–1027.
39. Vermimmen 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.
40. Wurtel H, Chartrand P (2006) Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res* 14: 477–495.
41. Amano T, Sagai T, Tanabe H, Mizushima 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.
42. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, et al. (2009) Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res* 19: 24–32.
43. Ferraiuolo MA, Rousseau M, Miyamoto C, Shenker S, Wang XQ, et al. (2010) The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res* 38: 7472–7484.
44. Kooren J, Palstra RJ, Klous P, Splinter E, von Lindern M, et al. (2007) Beta-globin active chromatin Hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. *J Biol Chem* 282: 16544–16552.
45. Kurukuti S, Tiwari VK, Tavosoidana G, Pugacheva E, Murrell A, et al. (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc Natl Acad Sci U S A* 103: 10684–10689.
46. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, et al. (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132: 422–433.
47. Nativo R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, et al. (2009) Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet* 5: e1000739.
48. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, et al. (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460: 410–413.
49. Hou C, Dale R, Dean A (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc Natl Acad Sci U S A* 107: 3651–3656.
50. Faure AJ, Schmidt D, Watt S, Schwale PC, Wilson MD, et al. (2012) Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Res*.
51. Conaway RC, Conaway JW (2011) Function and regulation of the Mediator complex. *Curr Opin Genet Dev* 21: 225–230.
52. Cuylen S, Haering CH (2010) A new cohesive team to mediate DNA looping. *Cell Stem Cell* 7: 424–426.
53. Oram SH, Thoms JA, Pridans C, Janes ME, Kinston SJ, et al. (2010) A previously unrecognized promoter of LMO2 forms part of a transcriptional regulatory circuit mediating LMO2 expression in a subset of T-acute lymphoblastic leukaemia patients. *Oncogene* 29: 5796–5808.
54. Landry JR, Bonadies N, Kinston S, Knezevic K, Wilson NK, et al. (2009) Expression of the leukemia oncogene Lmo2 is controlled by an array of tissue-specific elements dispersed over 100 kb and bound by Tal1/Lmo2, Ets, and Gata factors. *Blood* 113: 5783–5792.
55. Crable SC, Anderson KP (2003) A PAR domain transcription factor is involved in the expression from a hematopoietic-specific promoter for the human LMO2 gene. *Blood* 101: 4757–4764.
56. Royer-Pokora B, Rogers M, Zhu TH, Schneider S, Loos U, et al. (1995) The TTG-2/RBTN2 T cell oncogene encodes two alternative transcripts from two promoters: the distal promoter is removed by most 11p13 translocations in acute T cell leukaemia's (T-ALL). *Oncogene* 10: 1353–1360.
57. Soler E, Andrieu-Soler C, de Boer E, Bryne JC, Thongjuea S, et al. (2010) The genome-wide dynamics of the binding of Ldb1 complexes during erythroid differentiation. *Genes Dev* 24: 277–289.
58. Cheng Y, Wu W, Kumar SA, Yu D, Deng W, et al. (2009) Erythroid GATA1 function revealed by genome-wide analysis of transcription factor occupancy, histone modifications, and mRNA expression. *Genome Res* 19: 2172–2184.

59. Tallack MR, Whittington T, Yuen WS, Wainwright EN, Keys JR, et al. (2010) A global role for KLF1 in erythropoiesis revealed by ChIP-seq in primary erythroid cells. *Genome Res* 20: 1052–1063.
60. Wu C (1980) The 5' ends of *Drosophila* heat shock genes in chromatin are hypersensitive to DNase I. *Nature* 286: 854–860.
61. Myers RM, Stamatoyannopoulos J, Snyder M, Dunham I, Hardison RC, et al. (2011) A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS Biol* 9: e1001046.
62. Dickerman HW, Cheng TC, Kazazian HH Jr, Spivak JL (1976) The erythropoietic mouse spleen—a model system of development. *Arch Biochem Biophys* 177: 1–9.
63. Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, et al. (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* 36: 1065–1071.
64. Wilson NK, Foster SD, Wang X, Knezevic K, Schutte J, et al. (2010) Combinatorial transcriptional control in blood stem/progenitor cells: genome-wide analysis of ten major transcriptional regulators. *Cell Stem Cell* 7: 532–544.
65. Miles J, Mitchell JA, Chakalova L, Goyenechea B, Osborne CS, et al. (2007) Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus. *PLoS One* 2: e630.
66. Gribnau J, Diderich K, Pruzina S, Calzolari R, Fraser P (2000) Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol Cell* 5: 377–386.
67. Tuan D, Kong S, Hu K (1992) Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc Natl Acad Sci U S A* 89: 11219–11223.
68. Kim TK, Hemberg M, Gray JM, Costa AM, Bear DM, et al. (2010) Widespread transcription at neuronal activity-regulated enhancers. *Nature* 465: 182–187.
69. Pilon AM, Ajay SS, Kumar SA, Steiner LA, Cherukuri PF, et al. (2011) Genome-wide ChIP-Seq reveals a dramatic shift in the binding of the transcription factor erythroid Kruppel-like factor during erythrocyte differentiation. *Blood*.
70. Wijgerde M, Grosveld F, Fraser P (1995) Transcription complex stability and chromatin dynamics in vivo. *Nature* 377: 209–213.
71. 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.
72. Vakoc CR, Letting DL, Gheldof N, Sawado T, Bender MA, et al. (2005) Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol Cell* 17: 453–462.
73. Grill B, Wilson GM, Zhang KX, Wang B, Doyonnas R, et al. (2004) Activation/division of lymphocytes results in increased levels of cytoplasmic activation/proliferation-associated protein-1: prototype of a new family of proteins. *J Immunol* 172: 2389–2400.
74. Dekker J, Rippe K, Dekker M, Kleckner N (2002) Capturing chromosome conformation. *Science* 295: 1306–1311.
75. Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, et al. (2011) NCBI GEO: archive for functional genomics data sets—10 years on. *Nucleic Acids Res* 39: D1005–1010.
76. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10: R25.
77. Jothi R, Cuddapah S, Barski A, Cui K, Zhao K (2008) Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data. *Nucleic Acids Res* 36: 5221–5231.
78. Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, et al. (2010) The UCSC Genome Browser database: update 2010. *Nucleic Acids Res* 38: D613–619.