Published in final edited form as:
Nat Cell Biol. ; 13(8): 944–951. doi:10.1038/ncb2278.

Variegated gene expression caused by cell-specific long-range DNA interactions

Daan Noordermeer[1,4,5], **Elzo de Wit**[1,5], **Petra Klous**[1,5], **Harmen van de Werken**[1], **Marieke Simonis**[1], **Melissa Lopez-Jones**[2], **Bert Eussen**[3], **Annelies de Klein**[3], **Robert H. Singer**[2], and **Wouter de Laat**[1,6]

[1]Hubrecht Institute-KNAW and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands [2]Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, New York 10461, USA [3]Department of Clinical Genetics, Erasmus Medical Centre, PO Box 2040, 3000 CA Rotterdam, The Netherlands

Abstract

Mammalian genomes contain numerous regulatory DNA sites with unknown target genes. We used mice with an extra β-globin locus control region (LCR) to investigate how a regulator searches the genome for target genes. We find that the LCR samples a restricted nuclear subvolume, wherein it preferentially contacts genes controlled by shared transcription factors. No contacted gene is detectably upregulated except for endogenous β-globin genes located on another chromosome. This demonstrates genetically that mammalian trans activation is possible, but suggests that it will be rare. Trans activation occurs not pan-cellularly, but in 'jackpot' cells enriched for the interchromosomal interaction. Therefore, cell-specific long-range DNA contacts can cause variegated expression.

High-resolution profiling of transcription-factor binding sites, the discovery of conserved genetic elements and identification of regulatory sites using technologies such as DNaseI hypersensitive site mapping[1], has demonstrated that the number of genomic sites with transcription regulatory potential far exceeds the number of genes. One of the main challenges of the post-genomic era therefore is to assign function to each element. This requires understanding of the capacity of regulatory sites to reach over distance and identify specific target genes at the single-cell level. It is known that mammalian DNA elements can modulate the activity of distant genes on the same chromosome, up to a genomic distance of over a megabase[2,3]. The three-dimensional structure of the mammalian genome facilitates

[6]Correspondence should be addressed to W.d.L., w.delaat@hubrecht.eu.
[4]Present address: Laboratory of Developmental Genomics, School of Life Sciences, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland.
[5]These authors contributed equally to this work.

Note: Supplementary Information is available on the Nature Cell Biology website

AUTHOR CONTRIBUTIONS

D.N. and W.d.L. designed the experiments, analysed the data and, with help from E.d.W., wrote the manuscript. D.N. and P.K. carried out experiments. E.d.W. analysed 4C data and developed the automated FISH image analysis. H.v.d.W. analysed 4C and microarray expression data. M.S. carried out 4C experiments. M.L-J. and R.H.S. designed and synthesized RNA-FISH probes. B.E. and A.d.K. helped with the FISH experiments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

long-range gene regulation. This was first shown for the β-globin locus, which contains multiple β-globin genes arranged in the order of their developmental expression (Fig. 1a). Upstream of the genes resides a large regulatory DNA element that enhances expression of the β-globin genes up to ~100-fold[4]. The element is called a locus control region (LCR), as it has the capacity to confer position-independent and copy-number-dependent expression on reporter genes in transgenic assays in addition to its classical enhancer activity[5]. Otherwise, the β-globin LCR is not different from classical enhancers: it upregulates gene expression over distance, it functions in a tissue-specific manner and genes compete for its activity[6,7]. At its endogenous location, the LCR enhances expression of the β-globin genes through physical interactions, thereby looping out the intervening DNA that may carry non-responding genes[8–10] (Supplementary Fig. S1). A similar mode of action involving chromatin looping has been shown for enhancers in several other gene loci[11,12].

At a higher-order level of organization, microscopy studies have shown that genes can occupy differential positions in the nucleus depending on their expression status[13–15] and that regulatory DNA elements are instrumental in targeting these genomic regions to specific nuclear positions[16–19]. Moreover, recent observations suggest that functionally related genes on the same and different chromosomes may come together in the nuclear space for co-transcription[20]. Finally, some reports have suggested functional communication between regulatory sites located on different chromosomes[21–25]. Collectively, these studies raise questions on how genes and regulatory sequences explore the nuclear space to engage in functional crosstalk with preferred genomic partners.

RESULTS

Exploring mammalian transvection

To investigate the ability of regulatory DNA elements to relocate chromosomal regions in the nuclear interior and search for preferred target genes, we used transgenic mice with the human erythroid-specific β-globin LCR (hLCR) site-specifically integrated into a gene-dense region on chromosome 8. This region, 8C3–C4, contains many housekeeping genes[18]. Two transgenic lines were available, LCR-S and LCR-AS, each carrying the hLCR without globin genes at 8C3–C4, but in opposite orientations (Fig. 1b). In a previous study we showed that in erythroid cells each LCR detectably upregulates several housekeeping genes that directly surround the integration site up to sixfold, with the most distal one being 150 kilobases (kb) away. Microscopy studies revealed that both LCRs positioned 8C3–C4 more often outside its own chromosome territory, raising the question of to where the LCR migrates its integration site.

We reasoned that, in the most extreme situation, the hLCR could ‘search’ for a natural target gene present anywhere in the genome, including on its homologous chromosome, as is seen in transvection. The term transvection was coined to describe transcriptional regulation across (paired) homologous chromosomes, a phenomenon mainly studied in *Drosophila*, where homologues frequently pair[26,27]. Pairing is generally not observed in mammals[28], but may occur in special instances where mono-allelic expression needs to be ensured[29–31]. As mammalian chromosomes intermingle extensively[32], individual chromosomal segments may indeed invade the territories of other, possibly homologous, chromosomes. We generated transgenic mice carrying a human Aγ-globin reporter gene in one or the other orientation at exactly the same location in 8C3–C4 (Fig. 1b). The Aγ-globin gene is a human fetal globin gene that, in mouse transgenics carrying a full human globin cluster including the hLCR, is most highly expressed between embryonic day 10.5 and 12.5 (ref. 33; Fig. 1a). Crossing the Aγ-globin reporter mice with transgenic mice carrying the hLCR on the homologue revealed no increase in Aγ-globin gene expression at any of the relevant developmental stages (Fig. 1c). Fluorescence *in situ* hybridization (FISH) experiments also did not reveal increased

interaction between the homologues (Fig. 1d,e). Therefore, our experimental system provides no evidence for mammalian transvection involving the hLCR at 8C3–C4. Although these data do not rule out mammalian transvection at other genomic locations, they do strongly suggest that the ectopic hLCR does not have unlimited freedom to search the nuclear interior for natural target genes.

LCR motion is limited by chromosomal context

To investigate in more detail the ability of the hLCR to actively determine its genomic environment, we applied chromosome conformation capture[34] on chip (4C) technology to E14.5 fetal livers of wild-type and homozygous LCR-AS mice (Supplementary Fig. S2). 4C captures and identifies spatially proximal DNA sequences to enable an unbiased scan for DNA elements interacting with a locus of choice[35]. Analysis of the 4C data revealed that the ratio of inter-over intrachromosomal captures increased in the LCR-AS mice (Supplementary Fig. S3), in agreement with the hLCR causing looping out from the chromosome territory (ref. 18). This might be a reflection of the LCR actively engaging in interactions with new interchromosomal partners. To identify specific DNA interactions, sliding-window algorithms were applied that scan the linear chromosome templates for significant clustering of independently captured sequences[36,37]. Using both conventional analysis and a newly developed high-resolution analysis, we identified a highly similar set of interacting regions for 8C3–C4 with or without an integrated hLCR, both on the same chromosome (in *cis*) and on different chromosomes (in *trans*) (Fig. 2a–c and Supplementary Figs S4, S5a). Extensive fluorescence *in situ* hybridization experiments on cryo-sections (cryo-FISH; ref. 32) validated the 4C data (Fig. 2d,e). Some regions in *trans* showed increased interaction frequencies with the hLCR, again in agreement with the LCR causing looping out of the chromosome territory. However, we did not find chromosomal regions that exclusively interacted with the hLCR. Altogether, this demonstrates that the hLCR at 8C3–C4 does not search the genome to contact new preferred interaction partners. Rather, the chromosomal context of 8C3–C4 dictates the nuclear space that can be explored by the integrated hLCR.

The LCR contacts GATA-1 and EKLF regulated genes

Whereas the overall genomic environment did not change, a quantitative comparison of 4C results showed that the hLCR at 8C3–C4 captured a subset of pre-existing interchromosomal interaction partners more efficiently. This was most obvious for the α-globin locus on chromosome 11, but also for the endogenous β-globin locus on chromosome 7 (Fig. 3a). Homology between the human and mouse LCRs is limited (<10%) and therefore not expected to underlie the latter contact. The two loci have in common that they are very highly expressed and carry genes controlled by EKLF and GATA-1, two transcription factors that also bind to the LCR. Ranking genomic regions on the basis of their difference in 4C signal revealed significant enrichment ($P < 0.01$, hypergeometric test) of EKLF (refs 38,39)- and GATA-1 (ref. 40)-regulated genes, as well as of highly expressed genes, among regions contacted more strongly by the LCR (Supplementary Table S1). The same conclusion was drawn from an analysis per microarray probe (Fig. 3b). Each category of genes seems to independently attract the LCR (Fig. 3c and Supplementary Fig. S5b). We conclude that the LCR at 8C3–C4 behaves like a ‘dog on a lead’: the chromosomal context of 8C3–C4 imposes important constraints on its freedom to move, but within the restricted nuclear subcompartment that it occupies the LCR preferentially contacts genes that are controlled by shared transcription factors.

Interchromosomal gene activation by the LCR

Notwithstanding its restricted ability to change the spatial environment of 8C3–C4, the ectopic LCR is involved in many long-range interactions in *cis* and *trans*. We investigated

whether any of the contacted genes were upregulated by the LCR. For this, we profiled the transcriptomes of wild-type and LCR-transgenic littermates using microarrays. A small number of genes was found to be upregulated more than twofold in transgenics. These were the previously characterized genes proximal to the LCR in *cis* (ref. 18), plus a single gene in *trans*, *Hbb-bh1*, that was confirmed to be upregulated by quantitative PCR with reverse transcription (RT–qPCR) (Fig. 4 and Supplementary Fig. S6a–d).

Interestingly, *Hbb-bh1*, or *βh1*, is one of the endogenous β-globin genes normally expressed at high, LCR-dependent levels earlier in development, in embryonic blood cells (Fig. 1a). In E14.5 fetal livers the gene is looped away from the endogenous LCR (Supplementary Fig. S1). However, three independent probe-sets on the Affymetrix microarray and two independent RT–qPCR primer-sets reveal that the gene is not fully silent, but expressed at basal levels in wild-type fetal livers (Fig. 4 and Supplementary Fig. S6). Importantly, increased *Hbb-bh1* expression was consistently found in transgenics carrying the LCR in either one of the orientations when compared with their wild-type littermates (Fig. 4a,b), but not when 8C3–C4 exclusively carried a *Neo* selection cassette (Fig. 4c). No other β- or α-globin gene detectably changed in expression in the presence of the ectopic LCR (but see below), as judged by microarray analysis and RT–qPCR strategies that measure gene expression across cell populations (Fig. 4d and Supplementary Fig. S6e).

Trans-contacts cause variegated expression

Given that the endogenous β-globin locus carrying *βh1* is among the chromosomal regions that are preferentially contacted by the ectopic LCR (Fig. 3a), we hypothesized that such interchromosomal LCR–*βh1* contacts drive increased *βh1* expression. DNA and RNA-FISH analysis showed that the interaction occurs in 5–10% of the cells, whereas higher-resolution cryo-FISH identified contacts between 2 and 3% of the alleles (Supplementary Table S2). Irrespectively, this raises the question of how these relatively rare interchromosomal contacts can account for the overall twofold increase in transcript levels measured across the entire cell population. We reasoned that interaction frequencies measured by FISH in fixed cells may reflect chromatin dynamics, such that over time a given interaction occurs in every cell. Alternatively, they reflect genome conformations that are cell specific and relatively stable after mitotic exit, implying that, in a given nucleus, genomic loci sample overlapping or non-overlapping nuclear subvolumes. If the latter were true, ‘jackpot’ cells should exist with accumulated *βh1* messenger RNA levels in combination with frequent interchromosomal LCR–*βh1* interactions. To investigate this we carried out RNA FISH using a mixture of probes visualizing both primary transcripts and accumulated mRNA (ref. 41; Fig. 5a,b). In wild-type E14.5 fetal liver cells, we failed to detect the *βh1* primary transcript signals that were visible in E10.5 embryonic blood cells (Supplementary Fig. S7), consistent with its marked drop in expression during development. Accumulated cytoplasmic *βh1* mRNA was seen in only a small percentage (3.5%) of cells. These cells also contained the highest adult β-globin transcript levels (*Hbb-b1*, or *β-major*) and we therefore assumed they represent the most differentiated erythroid cells in the liver that had most time to accumulate β-globin mRNA. Notably, in only one out of 50 of these cells was an interchromosomal interaction between 8C3–C4 and β-globin observed, which does not exceed the interaction frequency of 5.1% measured across the entire wild-type red-blood-cell population (Fig. 5c and Supplementary Table S2). On the other hand, in LCR transgenic fetal liver cells the percentage of cells with high *βh1* mRNA levels is increased to 7.0%. More importantly, in these transgenics 15/50 cells (30%) with high *βh1* levels showed an interchromosomal interaction between 8C3–C4 (with the LCR) and β-globin, a highly significant ($P = 6.6 \times 10^{-7}$, hypergeometric test) increase in interaction frequency when compared with the 8.5% measured across all red blood cells in the transgenics (Fig. 5c and Supplementary Fig. S8a,b). As a control, we analysed α-versus β-globin co-localization in

transgenic cells and found no correlation between this interchromosomal interaction and *βh1* expression levels (Supplementary Fig. S8c).

We next reversed the question and asked whether cells showing the interchromosomal LCR–*βh1* interaction also had increased *βh1* transcript levels in their cytoplasm as compared with other red blood cells in the transgenic fetal livers. For this, we developed automated image-analysis software (see Methods) and analysed thousands of cells with respect to their intensity of *βh1* RNA FISH signal in the cytoplasm. The analysis showed that in transgenics, but not in wild type, cells with the interchromosomal interaction between 8C3–C4 and the β-globin locus more often have high *βh1* levels than the other cells from the same tissue (Fig. 6a). Interestingly, albeit less pronounced, the same was found for *β-major* (Fig. 6b), the adult β-globin gene located next to *βh1* in the endogenous globin locus (Fig. 1a). Its natural extremely high expression level precluded uncovering an extra contribution from the few cells with this interchromosomal interaction by cell-population-based expression assays such as microarrays or RT–qPCR. However, our analysis at the single-cell level revealed that also this adult β-globin gene benefits from contacts with the extra copy of the LCR. *βh1* is looped away from the endogenous mouse LCR and as such may be available for contacts with the ectopic human LCR. *β-major* dynamically forms and breaks contacts with the endogenous LCR (ref. 7), possibly providing opportunity for the ectopic LCR to also engage in contacts and further boost transcription. The genes encoding α-globin did not benefit from contacts with the ectopic LCR, as judged from the single-cell analysis strategy (Fig. 6c). We conclude that our transgenic mice contain a unique population of cells that have increased levels of mRNA for β-globin because these genes on chromosome 7 are contacted and trans activated by the ectopic LCR on chromosome 8.

DISCUSSION

One of the main challenges of the post-genomic era is to assign function to genomic sites, many of which have regulatory potential. Clearly, this cannot be done without considering the dynamics and spatial configuration of the genome. Here, we uncover properties of nuclear organization that dictate the action of regulatory elements in nuclear space. Our findings should contribute to a working model of genome function. The results demonstrate that regulatory DNA elements can search for preferred interaction partners, which in the case of the LCR are genes controlled by shared transcription factors. The ability to roam the nucleus is however heavily constrained by the chromosomal context. We predict the same to be true for almost all genomic locations, although the degree of constraint may vary. The concept of chromosomal context heavily influencing a gene's specific nuclear location seems to contradict more deterministic models of nuclear organization, where functionally related genes are proposed to meet at dedicated nuclear sites[20]. We cannot exclude that the LCR would have a more notable effect when placed at other genomic locations, or that other enhancers exist that are better capable of repositioning chromosomes and forming specific interchromosomal interactions. We note however that very few, if any, regulatory elements have been described with such a strong influence on gene expression and chromatin organization as the β-globin LCR.

An important finding of this study was that the ectopic, orphan, LCR on chromosome 8 contacted many different genes in *cis* and in *trans*, including those sharing a similar set of regulatory proteins, but that no measurable effect on the expression of most of them was detected. This suggests that, in mammals, ultralong-range gene regulation within and between chromosomes will be rare, or at least difficult to measure in cell populations. Two endogenous β-globin genes on chromosome 7 were the exception, as they were both upregulated by the ectopic LCR in cells with the relevant interchromosomal interaction. As these are natural (mouse) target genes of the (human) LCR, promoter compatibility and

spatial proximity seem essential for transcription regulation over distance. Interestingly, a few examples exist of endogenous tissue-specific enhancers activating not only target genes but also non-target genes that happen to be in physical proximity to the enhancer[42,43]. Our results open the possibility that such bystander activation may be more common in the genome, but appreciable only in individual cells that have their genome folded such that an enhancer and gene happen to be within contacting distance.

Our data provide genetic evidence for classical enhancer activity between mammalian chromosomes, where the genetic addition or deletion of a regulatory DNA element on one chromosome causes increased or reduced expression of a physically interacting gene on another chromosome. As such, we provide formal *in vivo* evidence that mammalian regulatory sites do not need an intervening chromatin fibre to propagate activating signals to responding gene promoters elsewhere in the genome, but that spatial proximity, in combination with enhancer–promoter compatibility, is sufficient for gene activation. Interchromosomal interactions between mammalian regulatory sites and genes have been observed before, but genetic evidence for *trans* activation was lacking so far. For example, the alternatively expressed T_H1 and T_H2 cytokine loci were seen to come together before their activation in naive T cells. On differentiation to T-helper 1 or 2 cells, the interactions between these signature loci were lost and the respective genes turned on[25]. The functional consequences of this interaction seemed complex, however, and different from classical enhancer activity. The deletion of a regulatory element in the T_H2 locus caused a delay, rather than a reduction, in the expression of the T_H1 gene, and intriguingly this effect was measurable only in differentiating T_H1 cells that no longer showed the interchromosomal interaction. The interaction was proposed to prepare loci for proper expression during subsequent T-helper cell specification, an activity not previously assigned to regulatory sites[25]. Interchromosomal gene regulation by a single enhancer was suggested to control the expression of all ~1,200 olfactory receptor genes spread across the genome[23], but deletion of the enhancer demonstrated that the enhancer only affects genes in *cis*[44,45]. In another study the activation of human interferon beta (*IFN-β*) expression in response to viral infection was reported to coincide with interchromosomal interactions with three Alu repeat elements harbouring cryptic NF-κB sites[21]. Although transfection experiments with plasmids carrying these elements supported the idea that the DNA interactions boost *IFN-β* expression, formal evidence for interchromosomal enhancer activity awaits demonstration that the chromosomal deletion of one of these repeats causes a drop in *IFN-β* expression. Finally, the imprinting control region (ICR) of the *H19-Igf2* locus has been the subject of several studies on interchromosomal DNA interactions[22,24,46]. The data did not reveal *trans* activation and were not necessarily consistent, as each study identified different interchromosomal interactions with different functional outcomes, possibly owing to the use of different cell types and/or experimental approaches.

An interesting observation from our artificial system is that interchromosomal interactions can lead to variegated levels of accumulated transcripts in the individual cells. We propose to term the observed phenomenon that cell-specific long-range DNA interactions cause variable gene expression levels among otherwise identical cells ‘spatial effect variegation’ (SEV; Fig. 6d). Stochastic cell-to-cell variation in gene expression, or transcriptional noise, is common in cell populations[47]. Our data open the possibility that SEV may be one of the underlying mechanisms of transcriptional noise. In such a scenario, the nature of the interacting region will determine whether gene expression goes up or down in the corresponding cell. This is different from position effect variegation[48], where variable expression of ectopically placed genes is classically thought to be caused by repressive effects from directly surrounding chromatin. Future research should indicate if SEV is acting on endogenous genes. If so, it may provide specific cells within a larger population with a mechanism to make autonomous cell-fate decisions, without the need for external signalling.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturecellbiology

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank J. Marteijn and W. Vermeulen for providing *Rad23a* knockout material, Y. Oz for counting FISH slides, V. Buckle for providing BAC probes, J. van Rheenen and the Hubrecht Imaging Center for help with image analysis and E. Splinter and other members of the group for assistance. This work was financially supported by grants from the Dutch Scientific Organization (NWO) (91204082 and 935170621) and a European Research Council Starting Grant (209700, '4C') to W.d.L.

References

1. Boyle AP, et al. High-resolution mapping and characterization of open chromatin across the genome. Cell. 2008; 132:311–322. [PubMed: 18243105]
2. Amano T, et al. Chromosomal dynamics at the Shh locus: limb bud-specific differential regulation of competence and active transcription. Dev Cell. 2009; 16:47–57. [PubMed: 19097946]
3. Kleinjan DJ, Coutinho P. Cis-ruption mechanisms: disruption of cis-regulatory control as a cause of human genetic disease. Brief Funct Genomic Proteom. 2009; 8:317–332.
4. Epner E, et al. The β-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse β-globin locus. Mol Cell. 1998; 2:447–455. [PubMed: 9809066]
5. Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell. 1987; 51:975–985. [PubMed: 3690667]
6. Wasylyk B, Wasylyk C, Augereau P, Chambon P. The SV40 72 bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. Cell. 1983; 32:503–514. [PubMed: 6297796]
7. Wijgerde M, Grosveld F, Fraser P. Transcription complex stability and chromatin dynamics *in vivo*. Nature. 1995; 377:209–213. [PubMed: 7675106]
8. Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P. Long-range chromatin regulatory interactions *in vivo*. Nat Genet. 2002; 32:623–626. [PubMed: 12426570]
9. Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active β-globin locus. Mol Cell. 2002; 10:1453–1465. [PubMed: 12504019]
10. de Laat W, Grosveld F. Spatial organization of gene expression: the active chromatin hub. Chromosome Res. 2003; 11:447–459. [PubMed: 12971721]
11. Spilianakis CG, Flavell RA. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. Nat Immunol. 2004; 5:1017–1027. [PubMed: 15378057]
12. Vernimmen D, De Gobbi M, Sloane-Stanley JA, Wood WG, Higgs DR. Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. EMBO J. 2007; 26:2041–2051. [PubMed: 17380126]
13. Brown KE, et al. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell. 1997; 91:845–854. [PubMed: 9413993]
14. Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA. Nuclear re-organisation of the Hoxb complex during mouse embryonic development. Development. 2005; 132:2215–2223. [PubMed: 15829525]
15. Skok JA, et al. Nonequivalent nuclear location of immunoglobulin alleles in B lymphocytes. Nat Immunol. 2001; 2:848–854. [PubMed: 11526401]

16. Hewitt SL, et al. Association between the Igk and Igh immunoglobulin loci mediated by the 3′ Igk enhancer induces 'decontraction' of the Igh locus in pre-B cells. Nat Immunol. 2008; 9:396–404. [PubMed: 18297074]
17. Lundgren M, et al. Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. Cell. 2000; 103:733–743. [PubMed: 11114330]
18. Noordermeer D, et al. Transcription and chromatin organization of a housekeeping gene cluster containing an integrated β-globin locus control region. PLoS Genet. 2008; 4:e1000016. [PubMed: 18369441]
19. Ragoczy T, Bender MA, Telling A, Byron R, Groudine M. The locus control region is required for association of the murine β-globin locus with engaged transcription factories during erythroid maturation. Genes Dev. 2006; 20:1447–1457. [PubMed: 16705039]
20. Schoenfelder S, et al. Preferential associations between co-regulated genes reveal a transcriptional interactome in erythroid cells. Nat Genet. 2010; 42:53–61. [PubMed: 20010836]
21. Apostolou E, Thanos D. Virus infection induces NF-κB-dependent interchromosomal associations mediating monoallelic IFN-β gene expression. Cell. 2008; 134:85–96. [PubMed: 18614013]
22. Ling JQ, et al. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science. 2006; 312:269–272. [PubMed: 16614224]
23. Lomvardas S, et al. Interchromosomal interactions and olfactory receptor choice. Cell. 2006; 126:403–413. [PubMed: 16873069]
24. Sandhu KS, et al. Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. Genes Dev. 2009; 23:2598–2603. [PubMed: 19933149]
25. Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA. Interchromosomal associations between alternatively expressed loci. Nature. 2005; 435:637–645. [PubMed: 15880101]
26. Lewis EB. The theory and application of a new method of detecting chromosomal rearrangements in *Drosophila melanogaster*. Am Naturalist. 1954; 88:225–239.
27. Pirrotta V. Transvection and chromosomal trans-interaction effects. Biochim Biophys Acta. 1999; 1424:M1–8. [PubMed: 10456029]
28. Bolzer A, et al. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 2005; 3:826–842.
29. Bacher CP, et al. Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat Cell Biol. 2006; 8:293–299. [PubMed: 16434960]
30. Hewitt SL, et al. RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. Nat Immunol. 2009; 10:655–664. [PubMed: 19448632]
31. Xu N, Donohoe ME, Silva SS, Lee JT. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat Genet. 2007; 39:1390–1396. [PubMed: 17952071]
32. Branco MR, Pombo A. Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. 2006; 4:780–788.
33. Strouboulis J, Dillon N, Grosveld F. Developmental regulation of a complete 70-kb human β-globin locus in transgenic mice. Genes Dev. 1992; 6:1857–1864. [PubMed: 1383089]
34. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002; 295:1306–1311. [PubMed: 11847345]
35. Simonis M, et al. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat Genet. 2006; 38:1348–1354. [PubMed: 17033623]
36. de Wit E, Braunschweig U, Greil F, Bussemaker HJ, van Steensel B. Global chromatin domain organization of the *Drosophila* genome. PLoS Genet. 2008; 4:e1000045. [PubMed: 18369463]
37. Simonis M, Kooren J, de Laat W. An evaluation of 3C-based methods to capture DNA interactions. Nat Methods. 2007; 4:895–901. [PubMed: 17971780]
38. Drissen R, et al. The erythroid phenotype of EKLF-null mice: defects in hemoglobin metabolism and membrane stability. Mol Cell Biol. 2005; 25:5205–5214. [PubMed: 15923635]

39. Hodge D, et al. A global role for EKLF in definitive and primitive erythropoiesis. Blood. 2006; 107:3359–3370. [PubMed: 16380451]
40. Welch JJ, et al. Global regulation of erythroid gene expression by transcription factor GATA-1. Blood. 2004; 104:3136–3147. [PubMed: 15297311]
41. Femino AM, Fay FS, Fogarty K, Singer RH. Visualization of single RNA transcripts *in situ*. Science. 1998; 280:585–590. [PubMed: 9554849]
42. Cajiao I, Zhang A, Yoo EJ, Cooke NE, Liebhaber SA. Bystander gene activation by a locus control region. EMBO J. 2004; 23:3854–3863. [PubMed: 15359275]
43. Lower KM, et al. Adventitious changes in long-range gene expression caused by polymorphic structural variation and promoter competition. Proc Natl Acad Sci USA. 2009; 106:21771–21776. [PubMed: 19959666]
44. Fuss SH, Omura M, Mombaerts P. Local and cis effects of the H element on expression of odorant receptor genes in mouse. Cell. 2007; 130:373–384. [PubMed: 17662950]
45. Nishizumi H, Kumasaka K, Inoue N, Nakashima A, Sakano H. Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis. Proc Natl Acad Sci USA. 2007; 104:20067–20072. [PubMed: 18077433]
46. Zhao Z, et al. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat Genet. 2006; 38:1341–1347. [PubMed: 17033624]
47. Raj A, van Oudenaarden A. Nature, nurture, or chance: stochastic gene expression and its consequences. Cell. 2008; 135:216–226. [PubMed: 18957198]
48. Muller HJ. Types of visible variations induced by X-rays in. Drosophila J Genet. 1930; 22:299–334.
49. Solovei I, et al. Spatial preservation of nuclear chromatin architecture during three-dimensional fluorescence *in situ* hybridization (3D-FISH). Exp Cell Res. 2002; 276:10–23. [PubMed: 11978004]
50. Levsky JM, Shenoy SM, Pezo RC, Singer RH. Single-cell gene expression profiling. Science. 2002; 297:836–840. [PubMed: 12161654]
51. Pau G, Fuchs F, Sklyar O, Boutros M, Huber W. EBImage–an R package for image processing with applications to cellular phenotypes. Bioinformatics. 2010; 26:979–981. [PubMed: 20338898]
52. Kolountzakis MN, Kutulakos KN. Fast computation of the Euclidian distance maps for binary images. Inform Process Lett. 1992; 43:181–184.
53. Vincent L, Soille P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. Pattern Analysis and Machine Intelligence, IEEE Trans. 2002; 13:583–598.
54. Smyth GK, Speed T. Normalization of cDNA microarray data. Methods. 2003; 31:265–273. [PubMed: 14597310]

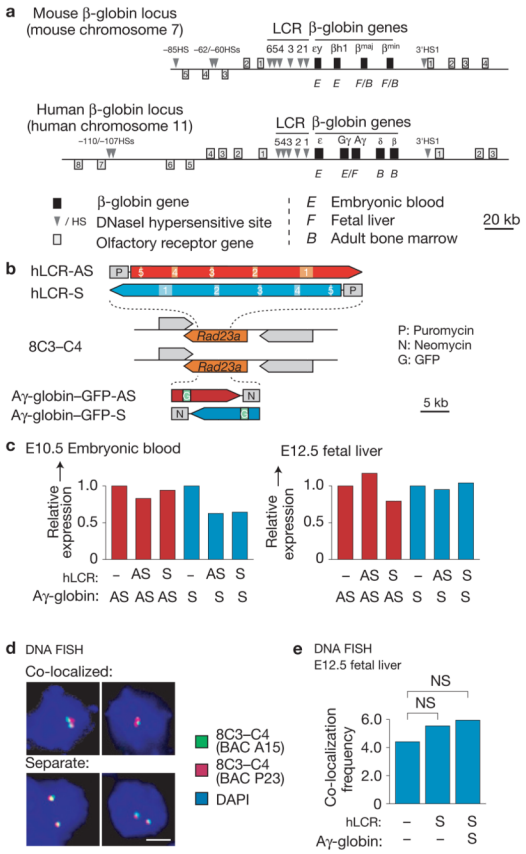

Figure 1.

An ectopic LCR does not activate a natural target gene on the homologous chromosome. (**a**) Schematic representation of the endogenous mouse and human β-globin loci. Below each globin gene, gene activity in (transgenic) mice is indicated. (**b**) Targeting strategy for the insertion of the human β-globin LCR and a human Aγ-globin–green fluorescent protein (GFP) reporter gene into the 8C3–C4 locus on mouse chromosome eight. (**c**) RT–qPCR of Aγ-globin transcript levels, normalized to *Hprt1* transcript levels. Data are from at least two independent samples. (**d**) Representative examples of DNA FISH showing co-localized and separate 8C3–C4 alleles. DAPI, 4,6-diamidino-2-phenylindole. Scale bar: 2 μm. (**e**) Co-localization frequencies of 8C3–C4 alleles. Significance levels are indicated above the graph (*G*-test).

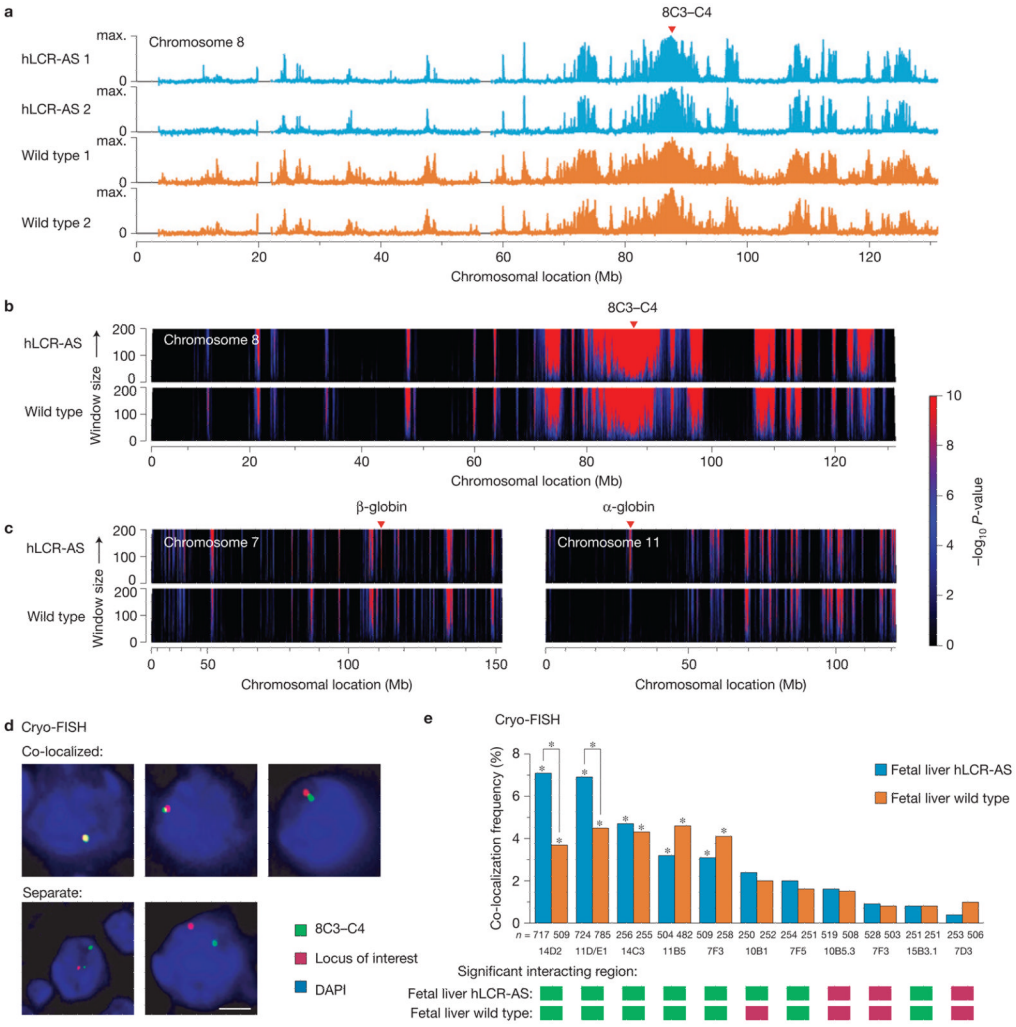

Figure 2.

Contacts of 8C3–C4 with and without the LCR are similar. (**a**) Intrachromosomal DNA interactions of 8C3–C4 with (top) and without (bottom) an integrated β-globin LCR are essentially similar, as determined by 4C analysed with a running-mean analysis of microarray data (average probe spacing: 7 kb). (**b**) Intrachromosomal DNA interactions of 8C3–C4 with (top) and without (bottom) an integrated β-globin LCR are essentially similar, as determined by 4C analysed with domainograms that visualize probability scores (*P*-values indicated with colour codes) for the clustering of positive 4C signals over windows ranging in size from 1 to 200 probes. (**c**) Interchromosomal 4C data for two chromosomes (7 and 11), analysed as described above. (**d**) Validation of 4C results by cryo-FISH; examples of results. Scale bar: 2 μm. (**e**) Interaction frequencies with a series of genomic regions, measured by cryo-FISH in wild-type and LCR transgenic fetal livers. The number of cells analysed (*n*) is indicated. Colour codes indicate the significance of the 4C signal (probe clustering), with green referring to $P < 0.01$ and red referring to $P \geq 0.01$.

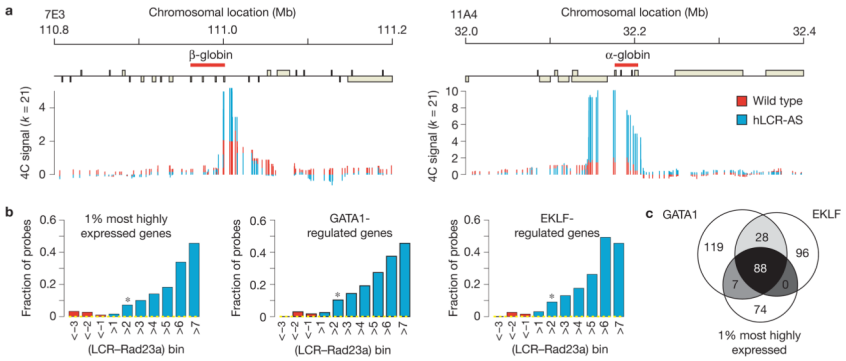

Figure 3.

Within a predetermined genomic environment the ectopic LCR shows preferential interactions with specific genes. (**a**) 4C results (running median over sliding windows of 21 probes) for 8C3–C4 with (blue) and without (red) the integrated LCR, at the endogenous β-globin locus on chromosome 7 (left) and the α-globin locus on chromosome 11 (right). (**b**) Probes were binned according to increasing difference in (LCR – wild-type) 4C signal and characterized depending on their location relative to highly expressed and GATA-1- and EKLF-regulated genes. The yellow dashed line represents the expected frequency on the basis of all probes. (**c**) Venn diagram showing the number of, and overlap between, probes captured more frequently in the integrated LCR 4C experiment for each category of genes analysed in the population marked with an asterisk in **b**.

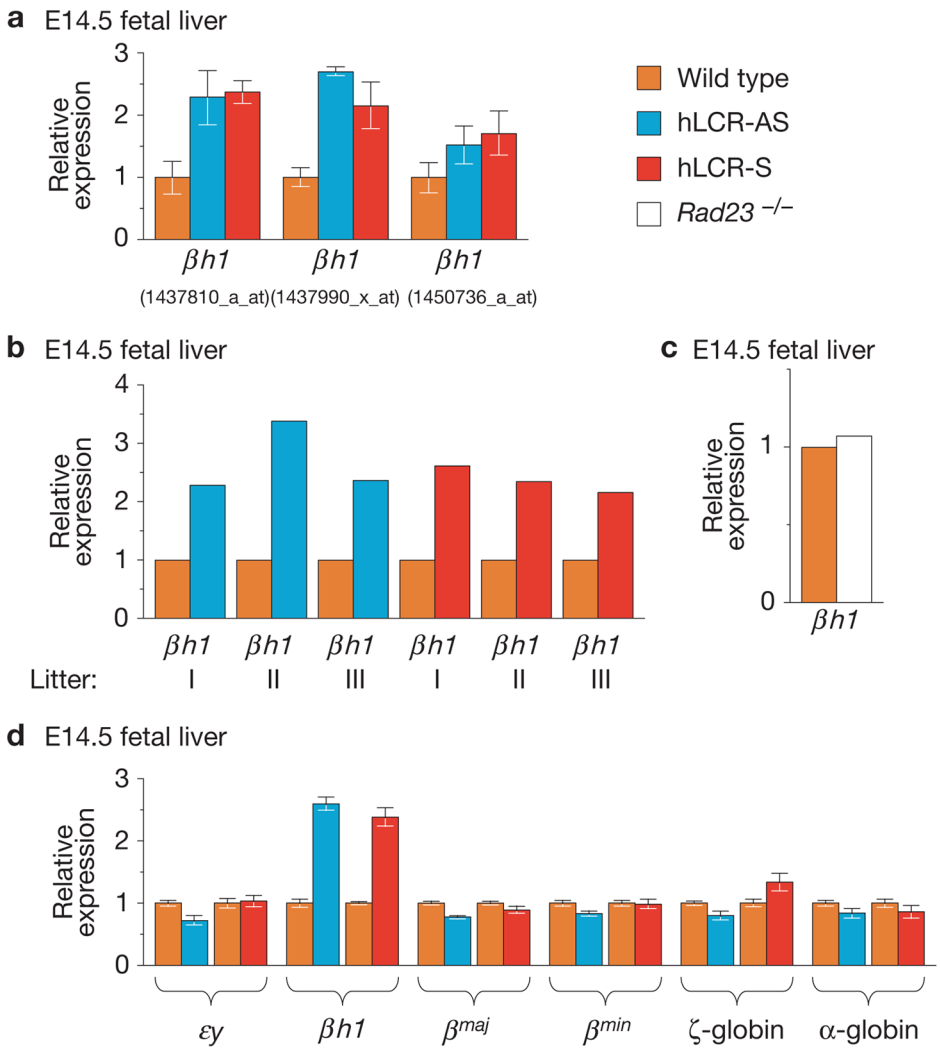

Figure 4.

The ectopic LCR on chromosome 8 enhances the expression of the endogenous *βh1* gene on chromosome 7. (**a**) Affymetrix gene-expression data for all probe-sets analysing *βh1* transcripts ($n = 3$). (**b**) RT–qPCR comparison of *βh1* gene expression between multiple wild-type and homozygous LCR-AS littermates, normalized to *Hprt1* transcript levels and to own wild-type littermates. (**c**) RT–qPCR analysis showing that the insertion at 8C3–C4 of a neomycin selection cassette instead of the hLCR does not lead to upregulation of the *βh1* gene. Data from two independent samples. (**d**) RT–qPCR analysis of expression of β- and α-globin genes in wild-type and homozygous LCR littermates. Error bars: standard error on the basis of 3 littermates (*n*).

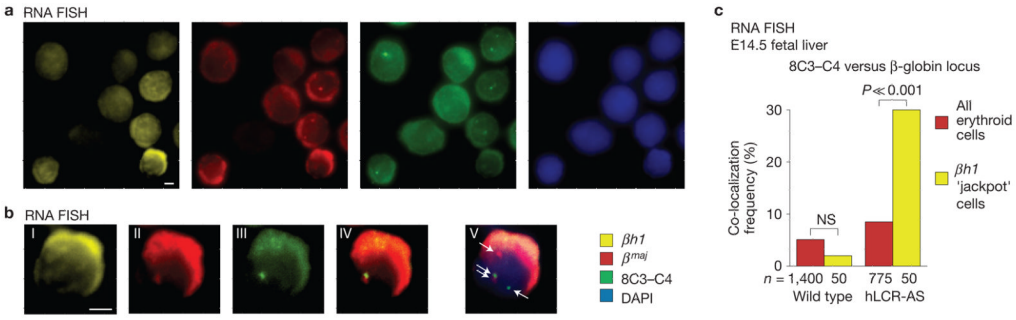

Figure 5.

Increased *βh1* mRNA levels in cells showing interchromosomal LCR–*βh1* interactions. (**a**) RNA FISH on E14.5 fetal liver cells, with one cell showing strongly increased *βh1* mRNA levels in the cytoplasm ('jackpot cell'). Scale bars: 2 μm. (**b**) Enlargement of the 'jackpot cell', revealing an interchromosomal interaction between the endogenous β-globin locus on chromosome 7 and the ectopic LCR on chromosome 8. (I–IV) Probes from one focal plane are shown separately and merged. (V) *Z* stack showing all RNA signals for β^{maj} and 8C3–C4. (**c**) Quantification of RNA FISH. Determining interchromosomal interaction frequencies between the endogenous β-globin locus and 8C3–C4 without (wild type) and with an integrated LCR (hLCR-AS), in all red blood cells and in '*βh1* jackpot cells'. The number of cells analysed (*n*) is indicated. NS: no significant difference.

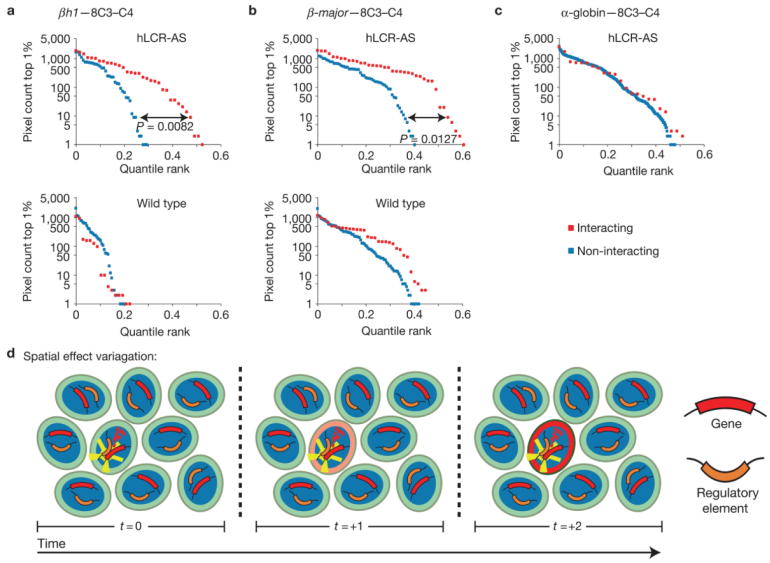

Figure 6.

Increased *βh1* and *β-major* mRNA levels in cells showing interchromosomal LCR–*βh1* interactions. (**a**,**b**) Automated RNA-FISH image-analysis (see Methods) results, showing that cells in which the ectopic LCR interacts in *trans* with the endogenous β-globin locus more often have high *βh1* (**a**) or *β-major* (**b**) transcript levels than cells that have the loci apart. (**c**) Cells in which the ectopic LCR interacts in *trans* with the endogenous α-globin locus do not differ in their levels of mRNA for α-globin from cells without this interchromosomal interaction. The probability score for the difference in distributions for interacting and non-interacting cells is calculated by a one-sided Kolmogorov–Smirnov test. (**d**) SEV: variegated expression among otherwise identical cells caused by cell-specific long-range DNA interactions (intra- or interchromosomal) that are relatively stable during interphase.