

## **A revised model for promoter competition based on multi-way chromatin interactions**

A. Marieke Oudelaar<sup>1,2\*</sup>, Caroline L. Harrold<sup>1\*</sup>, Lars L. P. Hanssen<sup>1</sup>, Jelena M. Telenius<sup>1,2</sup>, Douglas R. Higgs<sup>1</sup> and Jim R. Hughes<sup>1,2†</sup>

<sup>1</sup>MRC Molecular Haematology Unit, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.

<sup>2</sup>MRC WIMM Centre for Computational Biology, MRC Weatherall Institute of Molecular Medicine, Radcliffe Department of Medicine, University of Oxford, Oxford, UK.

\*These authors contributed equally: A. Marieke Oudelaar, Caroline Harrold.

†e-mail: [jim.hughes@imm.ox.ac.uk](mailto:jim.hughes@imm.ox.ac.uk)

## Abstract

Specific communication between gene promoters and enhancers is critical for accurate regulation of gene expression. However, it remains unclear how specific interactions between multiple regulatory elements and genes contained within a single chromatin domain are coordinated. Recent technological advances allow for the investigation of multi-way chromatin interactions at single alleles in individual nuclei. This can provide insights into how multiple regulatory elements cooperate or compete for transcriptional activation. We have used these techniques in a mouse model in which the  $\alpha$ -globin domain is extended to include several additional genes. This allows us to determine how the interactions of the  $\alpha$ -globin super-enhancer are distributed between multiple promoters in a single domain. Our data show that gene promoters do not form mutually exclusive interactions with the super-enhancer, but all interact simultaneously in a single complex. These findings show that promoters within the same domain do not structurally compete for interactions with enhancers, but form a regulatory hub structure, consistent with the recent model of transcriptional activation in phase-separated nuclear condensates.

## Introduction

An important question in current biology concerns the mechanisms by which genes are switched on and off during differentiation and development. Ultimately this is determined by interaction of the three fundamental regulatory elements of the genome: enhancers, promoters and boundary elements. The activity of these elements is closely related to the three-dimensional structure of the genome. Mammalian genomes are organized in topologically associating domains (TADs), which are self-interacting regions of chromatin, usually between 100 kb and 1 Mb in size (reviewed in Dekker and Heard, 2015; Dixon et al., 2016). The boundaries of TADs are often delineated by binding motifs for insulator proteins including CTCF-binding factor (CTCF) and the promoters of actively transcribed genes (Rao et al., 2014; Sanyal et al., 2012). There is increasing evidence that TADs are formed by a process of active extrusion of chromatin loops which is limited by these boundary elements (Fudenberg et al., 2018; 2016; Sanborn et al., 2015).

The specificity of interactions between regulatory elements is dependent on the TAD structure of the genome. Enhancers preferentially interact with gene promoters in the same TAD (Symmons et al., 2014) and disruption of TAD boundaries results in promiscuous enhancer-promoter interactions and disrupted gene activity (de Wit et al., 2015; Guo et al., 2015; Hanssen et al., 2017; Lupiáñez et al., 2015; Narendra et al., 2015). However, it is not clear how specificity between multiple enhancer elements and promoters contained within a single TAD is regulated. Enhancers often exert different effects on what appear to be equally accessible genes within individual TADs. It has been proposed that enhancer-driven transcription from different promoters within a TAD is dependent on distance, orientation or affinity of the enhancers with respect to the specific promoters (Furlong and Levine, 2018; Long et al., 2016). Previous studies have suggested that enhancers may only interact with one accessible promoter at a time. This has led to a model in which the pattern of gene expression within a TAD containing multiple genes is determined by competition between promoters for limited access to shared enhancers. Based on this model, it has been proposed that co-expression of multiple genes regulated by shared enhancers in a single TAD results from rapidly alternating interactions of these genes with the enhancers in a “flip-flop” mechanism (Bartman et al., 2016; Wijgerde et al., 1995).

Recent evidence suggests that transcriptional activation takes place in nuclear condensates, which contain a high concentration of transcription factors, coactivators and components of the basal transcription machinery recruited by enhancer elements (Boehning et al., 2018; Boija et al., 2018;

Cho et al., 2018b; Chong et al., 2018; Sabari et al., 2018). This implies that multiple enhancers within a TAD interact and function together in hub-like complexes, which have indeed been identified at a chromatin level (Allahyar et al., 2018; Oudelaar et al., 2018a). In the context of these recent findings, it is unclear if and how promoter competition occurs and what the underlying structural mechanism is.

We have recently developed Tri-C, a Chromosome Conformation Capture (3C) based approach which can analyze multi-way chromatin interactions at single alleles (Oudelaar et al., 2018a). Tri-C allows us to investigate whether promoters interact with enhancers in a mutually exclusive, one-to-one manner, or whether multiple promoters interact simultaneously with a shared set of enhancers in a hub-like structure. We have addressed this question using the well-characterized mouse  $\alpha$ -globin locus as a model. The  $\alpha$ -globin genes and their five enhancer elements, which fulfil the criteria for a super-enhancer (Hay et al., 2016), are located in a small TAD which is activated during erythroid differentiation (Brown et al., 2018). We have previously shown that *in vivo* deletion of two CTCF-binding sites at the upstream domain boundary results in an extension of the TAD and the incorporation of three upstream genes, which become highly upregulated under the influence of the strong  $\alpha$ -globin super-enhancer (Hanssen et al., 2017). These mutant mice provide an excellent model to analyze the interactions between genes co-regulated by a set of well-characterized enhancers in primary cells.

By performing Tri-C in erythroid cells in which the CTCF boundary is deleted, we show that the upregulated gene promoters preferentially interact in hub-like complexes containing both the  $\alpha$ -globin super-enhancer and the other active gene promoters in the domain. This shows that interactions between promoters and enhancers are not mutually exclusive and that there is no intrinsic structural competition between promoters for shared enhancers. These findings contribute to our understanding of the interplay between regulatory elements within and beyond TAD structures and the multiple layers of regulation that control gene expression.

## Results

We have previously defined all regulatory elements in and around the mouse  $\alpha$ -globin cluster (Brown et al., 2018; Hanssen et al., 2017; Hay et al., 2016) (Fig. 1). The duplicated  $\alpha$ -globin genes and the five globin enhancers (R1-R4 and Rm) lie within a small  $\sim$ 90 kb TAD. This TAD is flanked by predominantly convergent CTCF boundary elements. We have previously shown that deletion of the HS-38 and HS-39 CTCF-binding motifs causes strong upregulation of the upstream *Mpg*, *Rbbdf1* and *Snrnp25* genes in erythroid cells (Hanssen et al., 2017). To investigate how this deletion influences chromatin interactions with the  $\alpha$ -globin super-enhancer, we performed Capture-C from the viewpoint of the strongest enhancer element, R2, in primary erythroid cells derived from wild type (WT) mice and mice in which the CTCF-binding sites were deleted (D3839). This showed an extension of the interaction domain in the D3839 mice, causing increased interactions between the  $\alpha$ -globin enhancers and the *Mpg*, *Rbbdf1* and *Snrnp25* promoters (Fig. 1). The D3839 deletion thus creates an extended  $\sim$ 120 kb TAD in which the  $\alpha$ -globin super-enhancer upregulates multiple genes. This extended domain enables us to address the mechanism by which a super-enhancer interacts with multiple accessible gene promoters in a single TAD.

Although Capture-C produces high-resolution Chromosome Conformation Capture (3C) profiles (Davies et al., 2015), it predominantly generates pair-wise interaction data. It is therefore not possible to determine the higher-order structures in which the multiple promoters and enhancers in the extended  $\alpha$ -globin domain interact. Based on multi-way chromatin contacts generated by Tri-C, we have previously shown that the active  $\alpha$ -globin locus is organized in a hub structure, in which multiple elements of the  $\alpha$ -globin super-enhancer interact simultaneously with the  $\alpha$ -globin promoters in a regulatory complex (Oudelaar et al., 2018a). To examine this structure in the context of the extended  $\alpha$ -globin TAD containing multiple gene promoters, we performed a Tri-C experiment from the viewpoint of the R2 enhancer in primary D3839 and WT erythroid cells (Fig. 2 and Supplementary Fig. 2). Figure 2 shows the multi-way interactions detected by Tri-C. These are displayed in contact matrices in which we exclude the viewpoint of interest and plot the frequencies with which two elements interact simultaneously with this viewpoint in a single allele. Preferential, simultaneous interactions are visible as enrichments at the intersections between these elements, whereas mutually exclusive contacts between elements appear as depletions in the matrix. Consistent with our previous findings, we observe strong, simultaneous R2 interactions with the  $\alpha$ -globin promoter and enhancer elements in WT cells. These interactions are not

decreased in the D3839 cells, as would be expected if there was competition between these elements. Rather, there is a trend for increased interactions contributing to the  $\alpha$ -globin hub in the D3839 cells, though this is not significant (Fig. 2b-c; green). In addition to the multi-way contacts between the  $\alpha$ -globin enhancers and promoters, in D3839 we observe simultaneous interactions between the  $\alpha$ -globin enhancer elements and the *Mpg* and *Rhbdf1* promoters (Fig. 2b-c, purple). Interestingly, the R2 contact matrix also shows simultaneous contacts with both the  $\alpha$ -globin promoters and the *Mpg* and *Rhbdf1* promoters (Fig. 2b, grey). This indicates that all promoters in the extended D3839 TAD interact together in a single hub.

To allow more extensive examination of the simultaneous interactions occurring when the upstream genes interact with the  $\alpha$ -globin super-enhancer in D3839 cells, we next generated Tri-C data from the viewpoint of the *Mpg* promoter (Fig. 3 and Supplementary Fig. 3). Comparison of multi-way *Mpg* interactions in D3839 and WT cells reveals a strong increase in interactions downstream of *Mpg* after removal of the CTCF boundary. These interactions are strongest proximal to the *Mpg* promoter and reduce in strength beyond the R1 enhancer, which is located close to the HS-29 CTCF-binding site. We also observe a clear increase in more distal downstream multi-way interactions, predominantly with the regions containing the  $\alpha$ -globin promoters and enhancers. In a dynamic flip-flop model, the *Mpg* promoter would structurally compete with the  $\alpha$ -globin promoters for dynamic interactions with the  $\alpha$ -globin super-enhancer. Such mutually exclusive interactions would be reflected as a depletion of the corresponding multi-way interactions in the Tri-C matrix. However, we find preferential interactions between these elements. For example, when *Mpg* interacts with R1, it preferentially interacts with the  $\alpha$ -globin promoters (Fig. 3, green) and R4 enhancer (Supplementary Fig. 3). Similarly, we find enrichment of multi-way interactions between *Mpg*, R1 and the *Rhbdf1* promoter (Fig. 3, purple). This shows that *Mpg* preferentially interacts with the  $\alpha$ -globin super-enhancer in a complex that contains multiple enhancer elements and promoters. We also find that multi-way interactions between the three promoters are enriched (Fig. 3; orange), which further confirms that there is no structural competition between active promoters for contact with the super-enhancer within the extended TAD.

## Discussion

To investigate how multiple regulatory elements and genes contained within a single TAD structurally interact, we analyzed multi-way chromatin interactions in an engineered extended TAD containing the  $\alpha$ -globin super-enhancer and multiple gene promoters. We show that all gene promoters interact simultaneously with the five elements of the super-enhancer in a common regulatory hub. However, within the context of this extended TAD structure, the upstream non-globin genes do not interact as strongly and/or as frequently with the  $\alpha$ -globin super-enhancer compared to the  $\alpha$ -globin promoters. However, we show that when these genes form interactions with the  $\alpha$ -globin enhancers, they preferentially interact in a complex in which the  $\alpha$ -globin promoters are also present (Fig. 3, Supplementary Fig. 3). By comparing the  $\alpha$ -globin hub in WT and D3839 cells, we show that the inclusion of additional promoters to this complex does not weaken the interactions between the  $\alpha$ -globin promoters and enhancers and might even have an overall stabilizing effect on the hub (Fig. 2).

Our data thus show that multiple gene promoters can simultaneously interact with a shared super-enhancer at a single allele. This demonstrates that the previously reported “flip-flop” model of promoter competition, in which individual gene promoters interact with enhancers in a mutually exclusive manner, is not universally true and that there is no intrinsic competition between gene promoters for physical access to shared enhancers within a single TAD (Fig. 4). Our model is supported by recent live imaging experiments in *Drosophila* which showed coordinated bursting of two genes regulated by a single shared enhancer (Fukaya et al., 2016).

Our findings clarify how the activity of strong enhancers is distributed between the multiple genes surrounding these elements. In agreement with previous findings (de Wit et al., 2015; Guo et al., 2015; Hanssen et al., 2017; Lupiáñez et al., 2015; Narendra et al., 2015), enhancers and promoters do not interact beyond strong CTCF-binding sites at TAD boundaries, since removal of the HS-38/-39 boundary is required for the upstream genes to be activated by the  $\alpha$ -globin super-enhancer.

By contrast, within a single TAD, all promoters interact with the super-enhancer in a common nuclear compartment. This is consistent with the recent model of transcriptional activation in phase-separated nuclear condensates. However, even in the context of these cooperative

structures, the activity of enhancers may not always be distributed equally between all promoters in a TAD. This might partially be explained by the relative position of promoters with respect to the enhancers. Reported examples of promoter competition have often described situations where an active promoter located in between an enhancer and another, more distal promoter causes reduced activity of the distal promoter (Bartman et al., 2016; Cho et al., 2018a; De Gobbi et al., 2006; Wijgerde et al., 1995). It is possible that the proximal highly transcribed gene forms a barrier to loop extrusion (Brandão et al., 2019), due to accumulation of large amounts of transcriptional machinery and regulatory factors, which reduces interactions between the more distal promoter and the enhancers. However, the resulting functional competition does not occur due to mutually exclusive interactions, but because of reduced access to a cooperative condensate in which multiple regulatory elements can interact together.

Interestingly, we have shown that inclusion of the upstream genes in the  $\alpha$ -globin hub causes upregulation of their expression, but not to the exceptionally high levels of the downstream  $\alpha$ -globin genes (Hanssen et al., 2017). This could be explained by the lower frequency of interaction and inclusion of these genes in the chromatin hub, which likely corresponds to a nuclear condensate. However, it is also possible that biochemical processes within such condensates play a role, which might form another layer of regulation and potential competitive effects. With rapid technological advancements in single cell genomics and imaging, it will be interesting to relate chromatin structures to levels of gene expression in single cells in the future. This could provide important insights into the mechanisms that control gene activity and how mutations that disrupt regulatory chromatin structures contribute to human disease.



## Methods

### Animals and cells

We previously generated the D3839 mouse model, using TALEN and CRISPR-Cas9 to create small 19 bp and 26 bp deletions in the core CTCF-binding motifs of HS-38 and HS-39, respectively (Hanssen et al., 2017). We performed all described experiments in primary cells obtained from spleens of female D3839 or WT C57BL/6 mice treated with phenylhydrazine, and selected erythroid cells based on the erythroid marker Ter119 using magnetic-activated cell sorting, as previously described (Davies et al., 2015). Experimental procedures were in accordance with the European Union Directive 2010/63/EU and/or the UK Animals (Scientific Procedures) Act (1986) and protocols were approved through the Oxford University Local Ethical Review process.

### Capture-C – experimental procedure

We performed Capture-C experiments in 3 biological replicates of primary erythroid cells derived from WT and D3839 mice following the Next-Generation Capture-C protocol (Davies et al., 2015). We used the DpnII restriction enzyme for digestion during 3C library preparation. We designed capture oligonucleotides targeting the DpnII fragments containing the R1 and R2 enhancers using CapSequm (Hughes et al., 2014). Figure 1 shows the interaction profiles from the viewpoint of the R2 enhancer. Data from the R1 viewpoint have been published previously (GEO accession code GSE97871) (Hanssen et al., 2017). The Capture-C libraries were sequenced on the Illumina MiSeq platform (V2 chemistry; 150 bp paired-end reads).

### Capture-C – data analysis

We analyzed Capture-C data as previously described (Davies et al., 2015), using scripts available at <https://github.com/Hughes-Genome-Group/CCseqBasicS>. Because PCR duplicates are removed during data analysis, Capture-C accurately quantifies chromatin interactions (Oudelaar et al., 2017). The Capture-C profiles in Figure 1 represent the mean number of unique interactions per restriction fragment from 3 replicates, normalized for a total of 100,000 interactions on the chromosome analyzed, and scaled to 1,000. The differential profile highlights the interactions in D3839 cells after subtracting the normalized number of unique interactions in WT cells from those in D3839 cells. Interactions within a proximity zone of 1 kb around the viewpoint and with restriction fragments that were targeted by other capture oligonucleotides in the multiplexed capture procedure were excluded from analysis to prevent artefacts.

### Tri-C – experimental procedure

We performed Tri-C experiments in 3 biological replicates of primary erythroid cells derived from WT and D3839 mice following the protocol available on Protocol Exchange (Oudelaar et al., 2018b). We used the NlaIII restriction enzyme for digestion during 3C library preparation. We added Illumina TruSeq adaptors using NEBNext DNA Library Prep reagents and Ampure XP beads (Beckman Coulter: A63881). We prepared WT libraries using the NEBNext DNA Library Prep Master Mix Set for Illumina (New England Biolabs: E6040S/L) according to the manufacturer's protocol. For each biological replicate, we performed 2–3 parallel reactions using 6 ug 3C library for sonication and all recovered material (~4.5 ug) for the subsequent library preparation. We amplified each library preparation reaction with a different index, using 2 separate PCR reactions per reaction (a total of 4–6 PCR reactions per biological replicate) to maximize library complexity. This procedure resulted in a total of 7 technical replicates with unique indices. We prepared D3839 libraries using the NEBNext Ultra II DNA Library Prep Kit for Illumina (New England Biolabs: E7645S/L) according to the manufacturer's protocol. For each biological replicate, we sonicated 4 ug 3C library, after which we split all recovered material (~3 ug) over two parallel library preparation reactions. We amplified the two parallel reactions with the same index using 2 separate PCR reactions (a total of 4 PCR reactions per biological replicate) to maximize library complexity for each biological replicate. This resulted in a total of 3 replicates with unique indices. Because the Ultra II reagents are more efficient than the standard DNA Library Prep reagents, this resulted in comparable complexity for each biological replicate and similar data depth for both conditions (Supplementary Figs. 1 and 2). We pooled all libraries to enrich for our viewpoints of interest using a multiplexed double capture procedure with custom designed capture oligonucleotides (Supplementary Tables 1 and 2). The Tri-C libraries were sequenced on the Illumina NextSeq platform (V2 chemistry; 150 bp paired-end reads).

### Tri-C – data analysis

We analyzed Tri-C data using scripts available at <https://github.com/Hughes-Genome-Group/CCseqBasicS> and <https://github.com/oudelaar/TriC>. Briefly, we used the CCseqBasic pipeline (flags: --CCversion CS5 --nla --sonicationSize 700 --wobblyEndBinWidth 6) to perform the initial fastq processing and aligning of the data, filter out spurious ligation events and PCR duplicates, and exclude interactions with restriction fragments that were targeted by other capture oligonucleotides in the multiplexed capture procedure. We used a custom script to select reads with 2 or more reporters to calculate multi-way interaction counts between reporter fragments for each viewpoint. We visualized these interactions in contact matrices at 1 kb resolution, after

normalizing for the total counts in each matrix and correcting for the number of restriction fragments present in each bin. We integrated this workflow in the CCseqBasic pipeline, which is available at <https://github.com/Hughes-Genome-Group/CCseqBasicS>. To allow for direct comparisons between WT and D3839 cells, we scaled all contact matrices to 100 normalized interactions per bin. We derived differential matrices to highlight the interactions specific for each condition after subtracting the normalized interactions in WT cells from those in D3839 cells or vice versa. We also generated regular pair-wise interaction profiles based on the total interaction counts. These ‘Tri-Capture-C’ profiles were derived as described above (Capture-C – data analysis). To calculate the enrichment of multi-way interactions between cis-regulatory elements of interest (highlighted in Figs. 2 and 3), we calculated the counts in the bins in a 2 kb radius surrounding the foci of interest in the matrix and expressed these counts as a percentage of the total number of counts in the matrix. To analyze the differences between the WT and D3839 replicates, we used unpaired, two-tailed t-tests. Statistical analyses were performed with Student’s two-tailed t-tests using GraphPad Prism software.

## Acknowledgements

We thank Matthew Gosden for technical advice and help with next-generation sequencing. This work was supported by Wellcome (Genomic Medicine and Statistics PhD Programme, references 105281/Z/14/Z and 109110/Z/15/Z; Chromosome and Developmental Biology PhD Programme, reference 099684/Z/12/Z; Wellcome Trust Strategic Award, reference 106130/Z/14/Z) and the Medical Research Council (MRC Core Funding and Project Grant, reference MR/N00969X/1). A.M.O. is funded by a Stevenson Junior Research Fellowship at University College, Oxford.

## References

- Allahyar, A., Vermeulen, C., Bouwman, B.A.M., Krijger, P.H.L., Verstegen, M.J.A.M., Geeven, G., van Kranenburg, M., Pieterse, M., Straver, R., Haarhuis, J.H.I., et al. (2018). Enhancer hubs and loop collisions identified from single-allele topologies. *Nat Genet* *295*, 1306.
- Bartman, C.R., Hsu, S.C., Hsiung, C.C.S., Raj, A., and Blobel, G.A. (2016). Enhancer Regulation of Transcriptional Bursting Parameters Revealed by Forced Chromatin Looping. *Molecular Cell* *62*, 237–247.
- Boehning, M., Dugast-Darzacq, C., Rankovic, M., Hansen, A.S., Yu, T., Marie-Nelly, H., McSwiggen, D.T., Kokic, G., Dailey, G.M., Cramer, P., et al. (2018). RNA polymerase II clustering through carboxy-terminal domain phase separation. *Nat. Struct. Mol. Biol.* *25*, 833–840.
- Boija, A., Klein, I.A., Sabari, B.R., Dall'Agnesse, A., Coffey, E.L., Zamudio, A.V., Li, C.H., Shrinivas, K., Manteiga, J.C., Hannett, N.M., et al. (2018). Transcription Factors Activate Genes through the Phase-Separation Capacity of Their Activation Domains. *Cell* *175*, 1842–1855.e16.
- Brandão, H.B., Wang, X., Paul, P., van den Berg, A.A., Rudner, D.Z., and Mirny, L.A. (2019). RNA polymerases as moving barriers to condensin loop extrusion. *bioRxiv* *20*, 2057–2057.
- Brown, J.M., Roberts, N.A., Graham, B., Waithe, D., Lagerholm, C., Telenius, J.M., De Ornellas, S., Oudelaar, A.M., Scott, C., Szczerbal, I., et al. (2018). A tissue-specific self-interacting chromatin domain forms independently of enhancer-promoter interactions. *Nature Communications* *9*, 376.
- Cho, S.W., Xu, J., Sun, R., Mumbach, M.R., Carter, A.C., Chen, Y.G., Yost, K.E., Kim, J., He, J., Nevins, S.A., et al. (2018a). Promoter of lncRNA Gene PVT1 Is a Tumor-Suppressor DNA Boundary Element. *Cell* *173*, 1398–1412.e22.
- Cho, W.-K., Spille, J.-H., Hecht, M., Lee, C., Li, C., Grube, V., and Cisse, I.I. (2018b). Mediator and RNA polymerase II clusters associate in transcription-dependent condensates. *Science* *361*, 412–415.
- Chong, S., Dugast-Darzacq, C., Liu, Z., Dong, P., Dailey, G.M., Cattoglio, C., Heckert, A., Banala, S., Lavis, L., Darzacq, X., et al. (2018). Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science* *361*, eaar2555.
- Davies, J.O.J., Telenius, J.M., McGowan, S.J., Roberts, N.A., Taylor, S., Higgs, D.R., and Hughes, J.R. (2015). Multiplexed analysis of chromosome conformation at vastly improved sensitivity. *Nat Meth* *86*, 1202–1210.
- De Gobbi, M., Viprakashit, V., Hughes, J.R., Fisher, C., Buckle, V.J., Ayyub, H., Gibbons, R.J., Vernimmen, D., Yoshinaga, Y., de Jong, P., et al. (2006). A regulatory SNP causes a human genetic disease by creating a new transcriptional promoter. *Science* *312*, 1215–1217.
- de Wit, E., Vos, E.S.M., Holwerda, S.J.B., Valdes-Quezada, C., Verstegen, M.J.A.M., Teunissen, H., Splinter, E., Wijchers, P.J., Krijger, P.H.L., and de Laat, W. (2015). CTCF Binding Polarity Determines Chromatin Looping. *Molecular Cell* *60*, 676–684.
- Dekker, J., and Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. *FEBS Lett.* *589*, 2877–2884.

Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. *Molecular Cell* *62*, 668–680.

Fudenberg, G., Abdennur, N., Imakaev, M., Goloborodko, A., and Mirny, L.A. (2018). Emerging Evidence of Chromosome Folding by Loop Extrusion. *Cold Spring Harbor Symposia on Quantitative Biology* 034710.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. *CellReports* *15*, 2038–2049.

Fukaya, T., Lim, B., and Levine, M. (2016). Enhancer Control of Transcriptional Bursting. *Cell* *166*, 358–368.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. *Science* *361*, 1341–1345.

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D.U., Jung, I., Wu, H., Zhai, Y., Tang, Y., et al. (2015). CRISPR Inversion of CTCF Sites Alters Genome Topology and Enhancer/Promoter Function. *Cell* *162*, 900–910.

Hanssen, L.L.P., Kassouf, M.T., Oudelaar, A.M., Biggs, D., Preece, C., Downes, D.J., Gosden, M., Sharpe, J.A., Sloane-Stanley, J.A., Hughes, J.R., et al. (2017). Tissue-specific CTCF-cohesin-mediated chromatin architecture delimits enhancer interactions and function in vivo. *Nature Cell Biology* *19*, 952–961.

Hay, D., Hughes, J.R., Babbs, C., Davies, J.O.J., Graham, B.J., Hanssen, L.L.P., Kassouf, M.T., Oudelaar, A.M., Sharpe, J.A., Suci, M.C., et al. (2016). Genetic dissection of the  $\alpha$ -globin super-enhancer in vivo. *Nat Genet* *48*, 895–903.

Hughes, J.R., Roberts, N., McGowan, S., Hay, D., Giannoulatou, E., Lynch, M., De Gobbi, M., Taylor, S., Gibbons, R., and Higgs, D.R. (2014). Analysis of hundreds of cis-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat Genet* *46*, 205–212.

Long, H.K., Prescott, S.L., and Wysocka, J. (2016). Ever-Changing Landscapes: Transcriptional Enhancers in Development and Evolution. *Cell* *167*, 1170–1187.

Lupiáñez, D.G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, J.M., Laxova, R., et al. (2015). Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell* *161*, 1012–1025.

Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O., and Reinberg, D. (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* *347*, 1017–1021.

Oudelaar, A.M., Davies, J.O.J., Downes, D.J., Higgs, D.R., and Hughes, J.R. (2017). Robust detection of chromosomal interactions from small numbers of cells using low-input Capture-C. *Nucleic Acids Research*.

Oudelaar, A.M., Davies, J.O.J., Hanssen, L.L.P., Telenius, J.M., Schwessinger, R., Liu, Y., Brown, J.M., Downes, D.J., Chiariello, A.M., Bianco, S., et al. (2018a). Single-allele chromatin interactions identify regulatory hubs in dynamic compartmentalized domains. *Nat Genet* *50*, 1744–1751.

Oudelaar, A.M., Downes, D.J., and Hughes, J.R. (2018b). Tri-C. Protocol Exchange.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D Map of the Human Genome at Kilobase Resolution Reveals Principles of Chromatin Looping. *Cell* *159*, 1665–1680.

Sabari, B.R., Dall'Agnesse, A., Boija, A., Klein, I.A., Coffey, E.L., Shrinivas, K., Abraham, B.J., Hannett, N.M., Zamudio, A.V., Manteiga, J.C., et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. *Science* *361*, eaar3958.

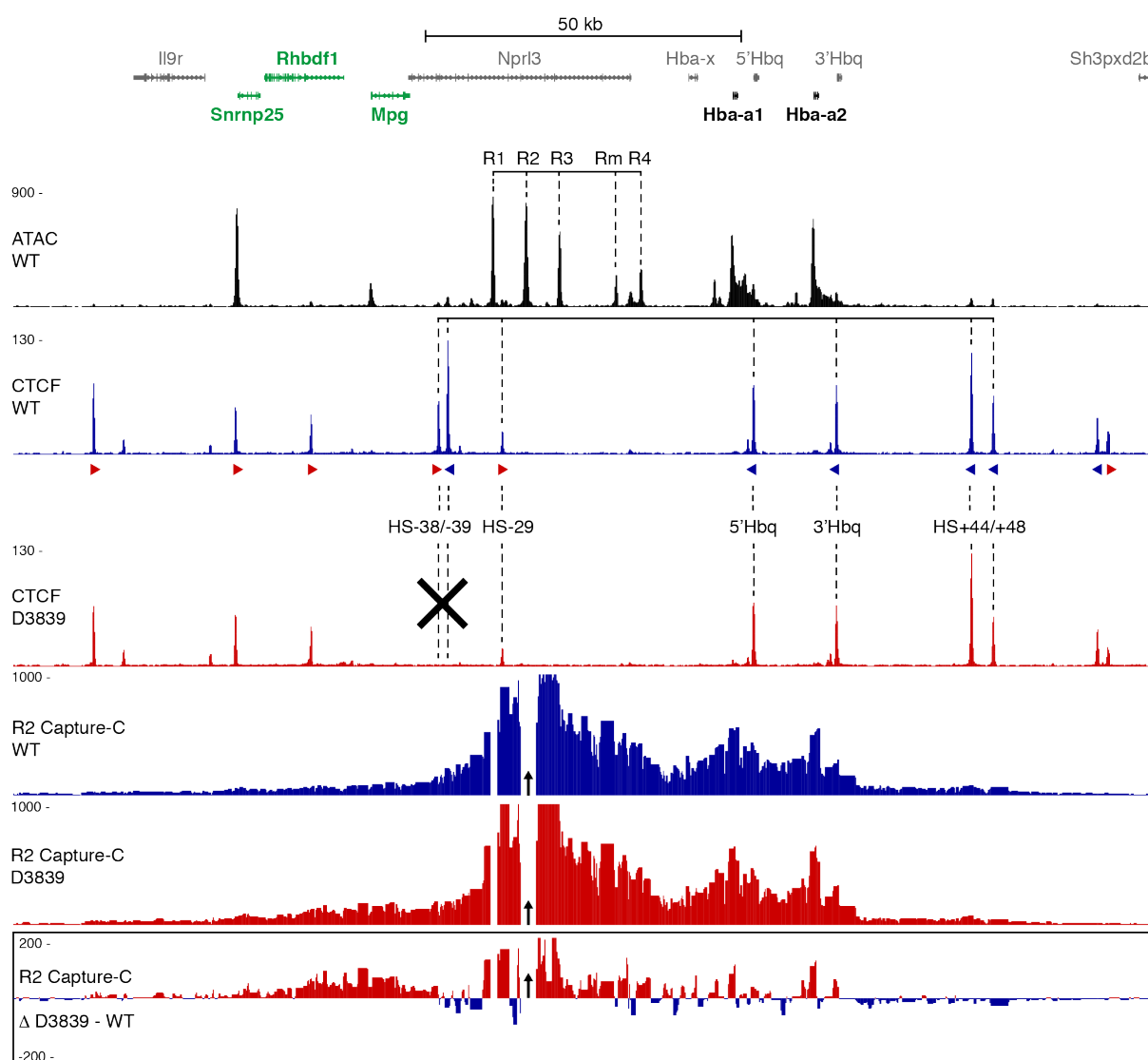
Sanborn, A.L., Rao, S.S.P., Huang, S.-C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proc. Natl. Acad. Sci. U.S.A.* *112*, E6456–E6465.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. *Nature* *489*, 109–113.

Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Eттwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. *Genome Research* *24*, 390–400.

Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. *Nature* *377*, 209–213.

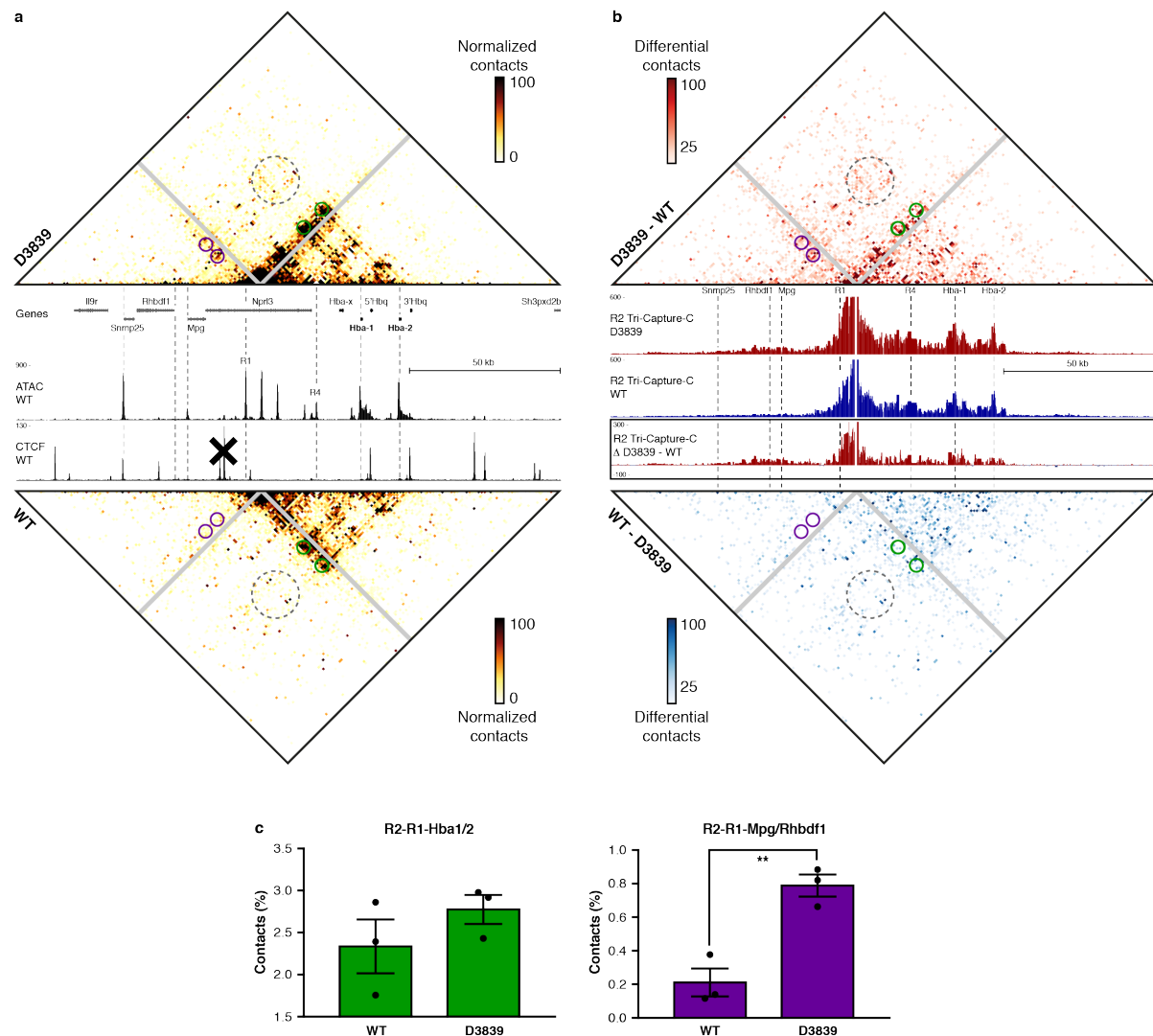
**Figure 1: Characterization of a CTCF boundary deletion upstream of the  $\alpha$ -globin locus.**



Gene annotation is shown at the top, with the  $\alpha$ -globin genes in bold and genes upregulated by the deletion highlighted in green. Open chromatin (ATAC in WT erythroid cells) is shown below, with the  $\alpha$ -globin enhancers highlighted. CTCF occupancy in WT (blue) and D3839 (red) erythroid cells is shown underneath, with the orientation of the CTCF binding motifs indicated by arrowheads (forward orientation in red; reverse orientation in blue). CTCF-binding sites of interested are highlighted and the deleted CTCF-binding sites are indicated with a black cross. The profiles below show Capture-C interactions from the viewpoint of the R2 enhancer (indicated with a black arrow) in WT (blue) and D3839 (red) erythroid cells, with a differential profile at the bottom. Profiles represent the mean number of normalized unique interaction counts per restriction fragment in 3 replicates. Coordinates (mm9): chr11:32,070,000–32,250,000.



**Figure 2: The formation of the enhancer-promoter hub at the  $\alpha$ -globin locus is not dependent on the flanking CTCF boundary.**

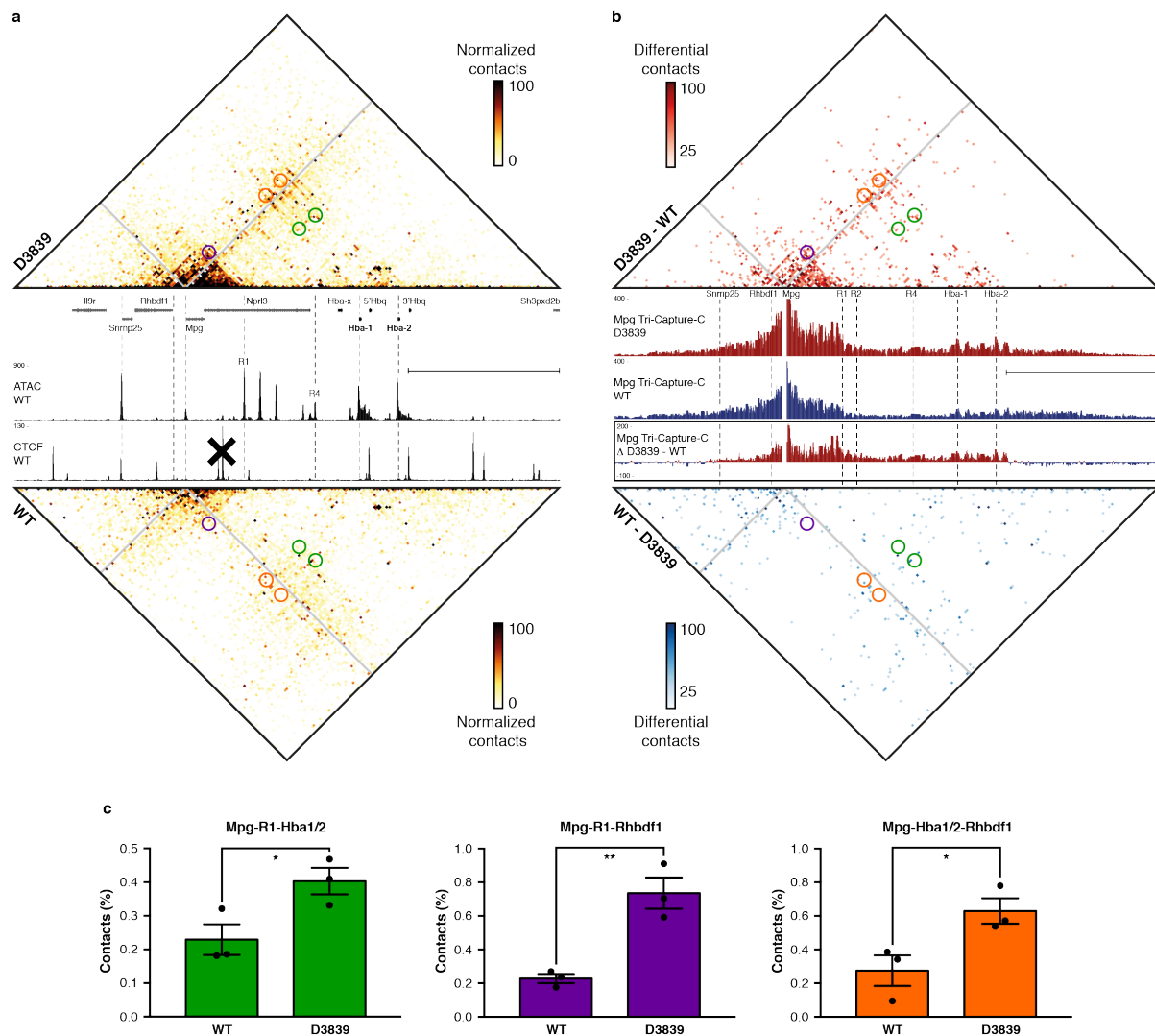


**(a)** Tri-C contact matrices showing multi-way chromatin interactions with R2 in D3839 (top) and WT (bottom) erythroid cells. Matrices represent mean numbers of normalized, unique contact counts at 1 kb resolution in 3 replicates with proximity contacts around the R2 viewpoint excluded (gray diagonal). Gene annotation, open chromatin (ATAC) and CTCF occupancy in WT erythroid cells are shown in the middle. Coordinates (mm9): chr11:32,070,000–32,250,000.

**(b)** Tri-C contact matrices showing differential multi-way chromatin interactions with R2 between D3839 and WT erythroid cells (top) and vice versa (bottom). Pair-wise interaction profiles derived from the Tri-C data from the R2 viewpoint (R2 Tri-Capture-C) are shown in the middle (D3839 in red, WT in blue), with a differential profile in the bottom panel. Coordinates (mm9): chr11:32,070,000–32,250,000.

**(c)** Quantification of multi-way contacts between R2, R1, and the  $\alpha$ -globin promoters (R2-R1-Hba1/2, green,  $P = 0.29$ ) and R2, R1, and the *Mpg* and *Rbdf1* promoters (R2-R1-*Mpg/Rbdf1*, purple,  $P = 0.0055$ ). Quantified contacts are highlighted with corresponding colors in the matrices above. Numbers represent the proportion of these 3-way contacts relative to the total in the matrix and are averages of 3 replicates, with individual data points overlaid as dot plots and the standard error of the mean denoted by the error bar.

**Figure 3: Deletion of a CTCF boundary results in the formation of a regulatory hub in which multiple gene promoters are incorporated.**

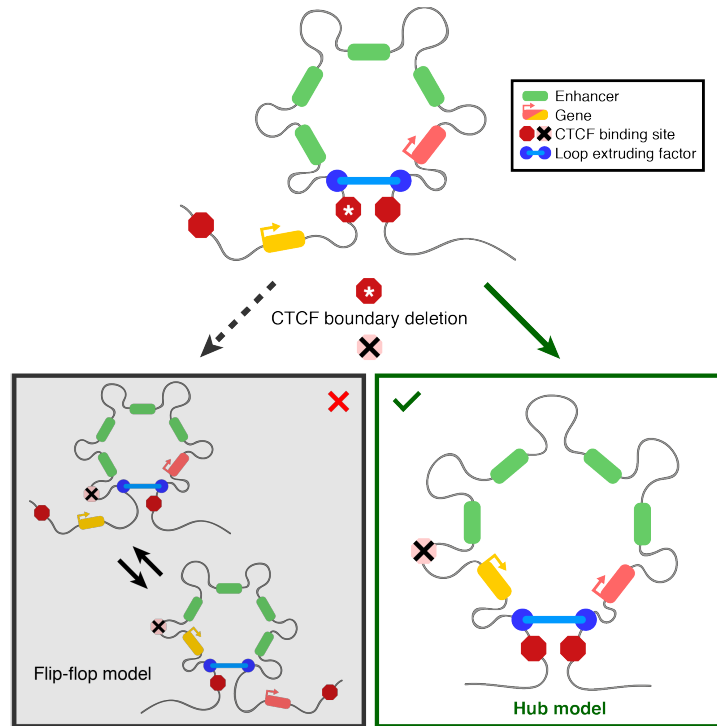


**(a)** Tri-C contact matrices showing multi-way chromatin interactions with *Mpg* in D3839 (top) and WT (bottom) erythroid cells. Matrices represent mean numbers of normalized, unique contact counts at 1 kb resolution in 3 replicates with proximity contacts around the *Mpg* viewpoint excluded (gray diagonal). Gene annotation, open chromatin (ATAC) and CTCF occupancy in WT erythroid cells are shown in the middle. Coordinates (mm9): chr11:32,070,000–32,250,000.

**(b)** Tri-C contact matrices showing differential multi-way chromatin interactions with *Mpg* between D3839 and WT erythroid cells (top) and vice versa (bottom). Pair-wise interaction profiles derived from the Tri-C data from the *Mpg* viewpoint (*Mpg* Tri-Capture-C) are shown in the middle (D3839 in red, WT in blue), with a differential profile in the bottom panel. Coordinates (mm9): chr11:32,070,000–32,250,000.

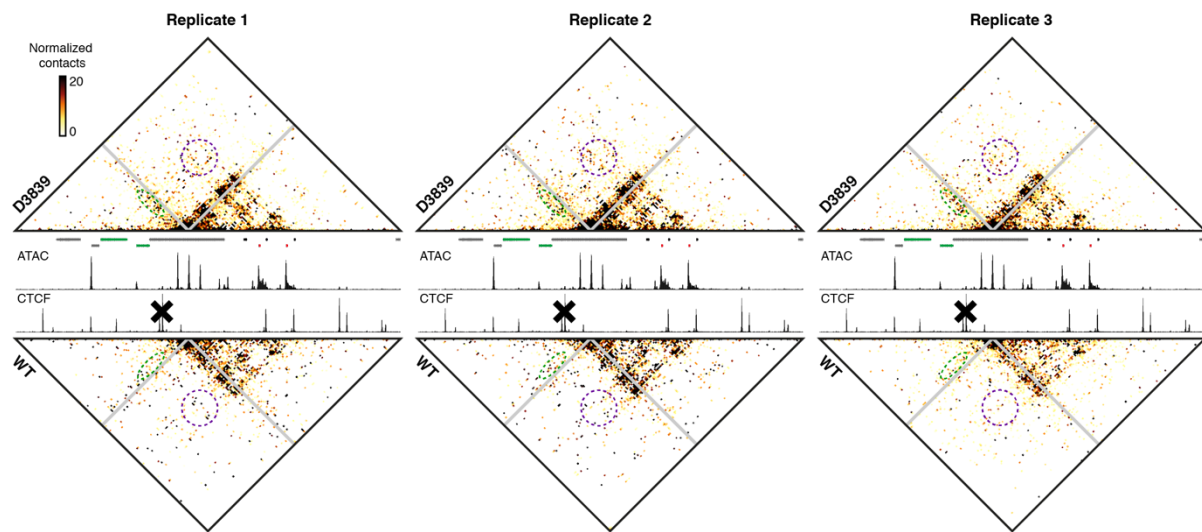
**(c)** Quantification of multi-way contacts between *Mpg*, R1 and the  $\alpha$ -globin promoters (*Mpg*-R1-Hba1/2, green,  $P = 0.046$ ); *Mpg*, R1 and the *Rbhd1* promoter (*Mpg*-R1-*Rbhd1*, purple,  $P = 0.0064$ ); and *Mpg*, the  $\alpha$ -globin promoters and the *Rbhd1* promoter (*Mpg*-Hba1/2-*Rbhd1*, orange,  $P = 0.040$ ). Quantified contacts are highlighted with corresponding colors in the matrices above. Numbers represent the proportion of these 3-way contacts relative to the total in the matrix and are averages of 3 replicates, with individual data points overlaid as dot plots and the standard error of the mean denoted by the error bar.

**Figure 4: Model of the structural interplay between the regulatory elements in the  $\alpha$ -globin locus upon removal of the upstream CTCF boundary.**



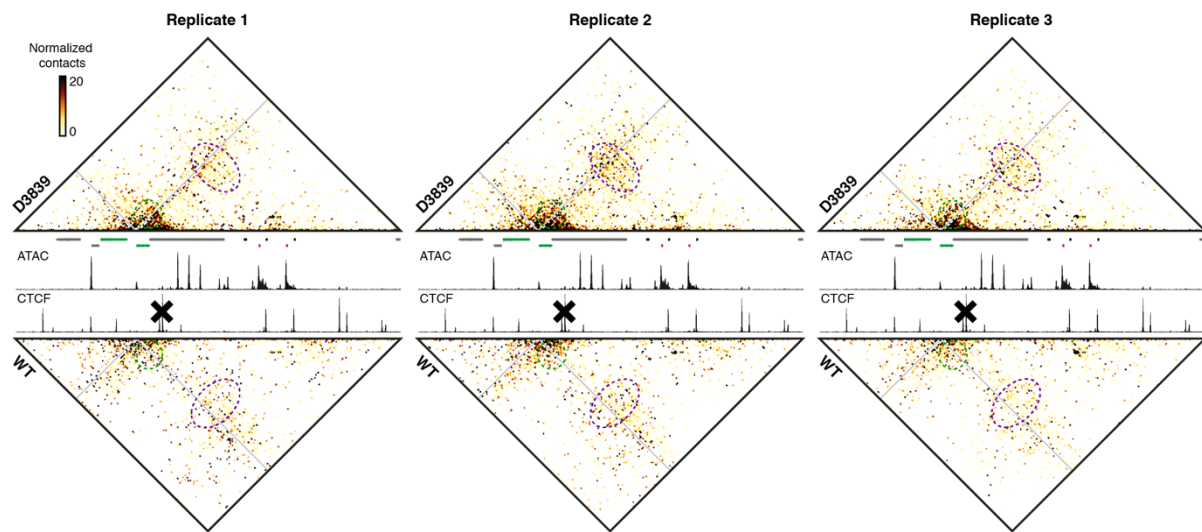
The CTCF boundary upstream of the  $\alpha$ -globin enhancers (green) normally constrains their activity to the downstream  $\alpha$ -globin genes (red). Removal of this boundary causes upregulation of the genes upstream of the  $\alpha$ -globin enhancers (yellow). Our data show that upregulation of these genes is not caused by dynamically switching interactions between the  $\alpha$ -globin enhancers and individual promoters (left), but by the formation of a regulatory hub in which all regulatory elements interact simultaneously (right).

### Supplementary Figure 1: Reproducibility of R2 Tri-C contact matrices.



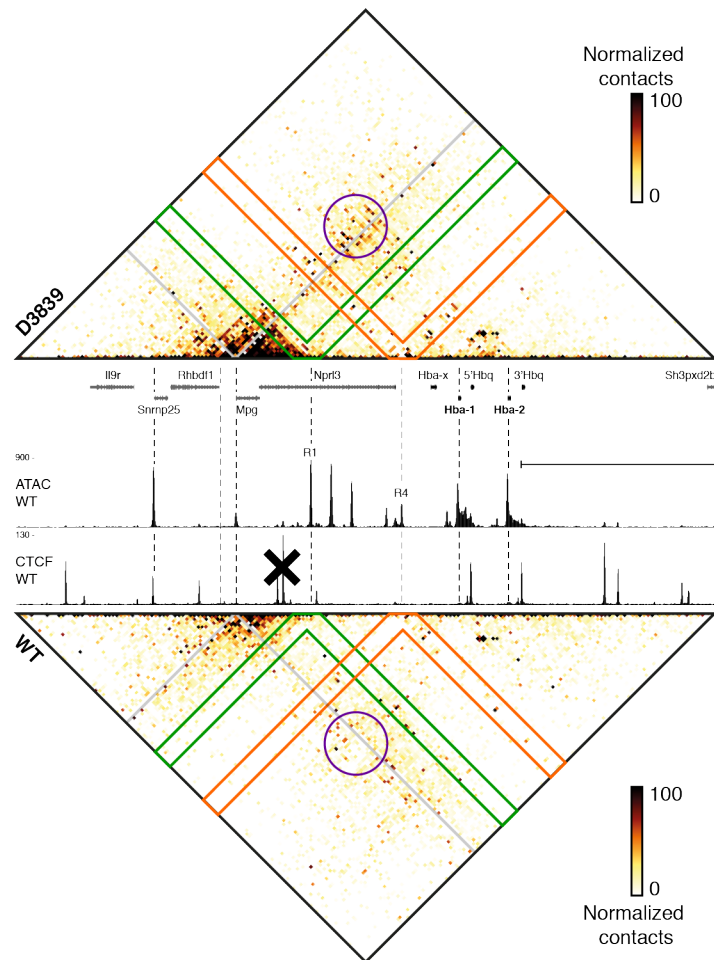
Tri-C contact matrices showing multi-way chromatin interactions with R2 in individual biological replicates of D3839 (top) and WT (bottom) erythroid cells. Matrices represent normalized, unique contact counts at 1 kb resolution with proximity contacts around the R2 viewpoint excluded (gray diagonal). Individual replicates show similar patterns of increased R2 interactions with R1 and the Mpg/Rhbdf1 promoters (green) and the Mpg/Rhbdf1 promoters and the  $\alpha$ -globin promoters (purple) in the D3839 cells compared to WT cells. Gene annotation, open chromatin (ATAC) and CTCF occupancy in WT erythroid cells are shown in the middle. Coordinates (mm9): chr11:32,070,000–32,250,000.

## Supplementary Figure 2: Reproducibility of Mpg Tri-C contact matrices.



Tri-C contact matrices showing multi-way chromatin interactions with Mpg in individual biological replicates of D3839 (top) and WT (bottom) erythroid cells. Matrices represent normalized, unique contact counts at 1 kb resolution with proximity contacts around the Mpg viewpoint excluded (gray diagonal). Individual replicates show similar patterns of increased proximal Mpg interactions, including R1 and the Mpg/Rhbdf1 promoters (green) and the Rhbdf1 promoter and the  $\alpha$ -globin enhancers/promoters (purple) in the D3839 cells compared to WT cells. Gene annotation, open chromatin (ATAC) and CTCF occupancy in WT erythroid cells are shown in the middle. Coordinates (mm9): chr11:32,070,000–32,250,000.

### Supplementary Figure 3: Multi-way interactions with the Mpg promoter.



Tri-C contact matrices showing multi-way chromatin interactions with Mpg in D3839 (top) and WT (bottom) erythroid cells. Matrices represent mean numbers of normalized, unique contact counts at 1 kb resolution in 3 replicates with proximity contacts around the Mpg viewpoint excluded (gray diagonal). Gene annotation, open chromatin (ATAC) and CTCF occupancy in WT erythroid cells are shown in the middle. Coordinates (mm9): chr11:32,070,000–32,250,000. To emphasize that the Mpg promoter preferentially interacts with the  $\alpha$ -globin enhancers in a complex which includes the  $\alpha$ -globin and Rhbdf1 promoters, we have highlighted the regions of the contact matrices that show all the multi-way interactions between Mpg and R1 (green) and between Mpg and R4 (orange). When Mpg interacts with R1 or R4 in D3839 cells, there are clear enrichments over the other  $\alpha$ -globin enhancers and the  $\alpha$ -globin and Rhbdf1 promoters, indicating that Mpg preferentially interacts with these elements in a complex. The formation of a structure in which multiple promoters interact together is also evident from the increased Mpg interactions with the  $\alpha$ -globin and Rhbdf1 promoters (purple) in D3839 cells.

**Supplementary Table 1. Tri-C viewpoints.** Overview of the coordinates and sizes of the restriction fragments used as viewpoints in the Tri-C experiments. The oligonucleotide pools we used for viewpoint enrichment also contained oligonucleotides targeting the following NlaIII fragments: chr11:32137188-32137324, chr11:32100062-32100217 and chr11:32160146-32160318. These restriction fragments were excluded from analysis to prevent artefacts.

	<b>NlaIII fragment coordinates</b>	<b>NlaIII fragment size (bp)</b>
<b>R2</b>	chr11:32,150,926-32,151,102	176
<b>Mpg</b>	chr11:32,126,462-32,126,692	230

**Supplementary Table 2. Tri-C capture oligonucleotides.** Overview of the sequences of the Tri-C capture oligonucleotides we used to enrich for viewpoints of interest. These 120 bp sequences were designed to target the middle of the restriction fragments listed in Supplementary Table 1.

	<b>Sequence</b>
<b>R2</b>	GGTCAAAGTAGCATACACCCATCTGGAACCTATCAGTGACCATAGTCAACAGCAGGTGTACACA CCCAGGCCAAGGGTGGAGCAGACCACTGTGGGATCTATGGAGATGCTTGAACGAGC
<b>Mpg</b>	TCCGGTGGCCTGGCCTGTGCTGGCGGCGACTAGATGCCCGCGCGCGGTGGTAGTGCGCGCCC GGGCAGAGGAGCCCTAAAACCGGTGTCCGTGACCCTGCTCCCCGACACCGAGCAGCCT